University College London

Institute of Structural and Molecular Biology

BIOCHEMICAL AND STRUCTURAL CHARACTERISATION OF PATHOLOGICAL ALPHA$_1$-ANTITRYPSIN POLYMERS

Submitted for the degree of Doctor of Philosophy

October 2018

Declaration

I, confirm that the work presented in this thesis is my own. Where information has been derived from other sources or carried out in collaboration with others, I confirm this has been explicitly indicated in the text.

Acknowledgements

I would like to give the biggest thanks to my UCL supervisors Professor David Lomas and Dr James Irving for all their help and support over the last 3 years. David's enthusiasm, positivity and patience has been greatly appreciated on this challenging project. James' advice and guidance has been invaluable in putting this thesis together and he has taught me so much, both in and out of the lab. I have really enjoyed working with him and will miss his quick and witty humour!

I'd also like to thank all the other members of the group for all the fun along the way: Dr Nina Heyer-Chauhan, Dr Riccardo Ronzoni, Dr Freddie Partridge, Magd Badaoui, Jing He, Scott Nanda, Josh Wort, Michael Poeschla, Mattia Laffranchi, Lathan Liou and particularly Professor David Sattelle for his supportive and hilarious emails. I thoroughly enjoyed the lab summer visits from Dr Elena Miranda and Dr Juan Perez, speaking Spanish and deciding on our favourite hybridomas! Not forgetting the two most important people: soon-to-be Dr Anwen Brown and soon-to-be Dr Alistair Jagger, my fellow PhD buddies. These two really made all the scientific 'failures' so much more bearable and have frequently made me laugh until I cried. I will miss our Tuesday Caffé Nero dates and Friday Birkbeck bar nights. I really couldn't have wished for better colleagues and have made some friends for life.

I am hugely grateful to my secondary supervisor at Birkbeck, Professor Elena Orlova for welcoming me to the world of electron microscopy. Elena is a fantastic teacher, with many amusing analogies for complex concepts, which were a great source of entertainment during our "silly questions" meetings. Thanks also to Dr Adam Redzej for training me on the microscopes and becoming a good friend and to Dr Natasha Lukoyanova for her support with the project. Dr Sarah Faull and Professor Bibek Gooptu must also be thanked for all their hard work in starting the electron microscopy work of alpha$_1$-antitrypsin polymers.

Special acknowledgements to Dr Andrew Gardner, my parents and friends who have been a constant support network throughout and have lived through all the ups and downs of a PhD with me. Lastly, I would like to thank the Wellcome Trust for funding my PhD and for its support throughout.

Abstract

The hereditary disease alpha$_1$-antitrypsin deficiency (AATD), most prevalent in those of Northern European ancestry, is a leading genetic cause of early-onset emphysema and liver cirrhosis. Currently, the only effective treatment is a transplant of the lungs or liver. The disease is caused by mutations in the gene encoding alpha$_1$-antitrypsin (AAT), which promote misfolding and aggregation into polymer chains. These polymers accumulate and induce cellular damage to hepatocytes, and the resultant lack of AAT in the circulation permits uncontrolled remodelling, and eventual breakdown, of lung tissue.

The mechanism of polymerisation and the linkage between individual subunits of the polymer found in affected individuals is unknown. Current structural models of polymerisation are based on artificially-induced *in vitro* polymers, but their relevance to *in vivo* systems is disputed. Polymers formed by heating are believed to represent polymers found in patients. A structural understanding of the mechanism by which AAT polymerises will be useful for the development and evaluation of therapeutic strategies.

Conformational-specific monoclonal antibodies and electron microscopy have been used in combination to study the structure of heat-induced and patient-derived polymers. Negative stain-electron microscopy (NS-EM) of polymers, with Fab fragments attached to each AAT subunit, has generated similar reconstructions of a dimer for both polymer types. This supports the notion that heat-induced polymers are a suitable model for *ex vivo* polymers. An initial cryo-EM reconstruction of the labelled *ex vivo* dimer and a single subunit has also been obtained, in preparation for a higher-resolution cryo-EM dataset. Fitting of the polymerisation models into the EM data best supports the C-terminal model of polymerisation. Lastly, a novel monoclonal antibody that is selective for AAT monomer over polymer, called 1D9, has been developed. X-ray crystallography reveals it binds to a region of AAT most exposed in the monomer; a characteristic which could mean it has future applications in diagnosing AATD.

Impact statement

There are no drugs to treat the disease alpha$_1$-antitrypsin deficiency (AATD). It results from mutations that lead to polymerisation of alpha$_1$-antitrypsin (AAT) in the liver. This severely reduces AAT levels in the circulation and lungs, leading to emphysema. I therefore investigated the polymerisation mechanism of AAT, i) structurally, using the technique of electron microscopy (EM) to study artificially-made and patient-derived AAT polymers and ii) biochemically, using monoclonal antibodies that have specific recognition profiles for different AAT conformations. Understanding the mechanism of AAT polymer formation may offer new therapeutic avenues.

The applications of the EM work are wide-ranging. For the AAT field, the structures of the AAT polymer have been useful in testing the currently proposed models of polymerisation. AAT polymerisation underlies the liver and lung disease of those with AATD and polymerisation of related proteins (serpins) can also cause thrombosis and a sub-type of dementia. It is hoped insights into the polymer structure will be applicable to all these diseases and inform highly sought-after future therapeutic approaches to stop polymerisation from occurring. On a broader academic level, this work addressed the use of *in vitro* models to represent what occurs *in vivo*. Here, polymers were uniquely extracted and studied from a human liver; previous research has focussed on artificially-made AAT polymers. To enable this, novel purification methodologies from the liver tissue were developed. The importance of studying patient-derived material for human disease has been highlighted and must be considered for other diseases such as Alzheimer's. Finally, cryo-EM of the AAT polymer will contribute to the wider cryo-EM community. The technique has significant technological promise, is attracting substantial pharmaceutical investment and scientists have recently been awarded the Nobel Prize for its development. The AAT polymer pushes the boundaries of what is currently technically possible to detect and structurally determine due to its small size and flexibility. It has been necessary to develop new methods of image analysis and workflow. At conferences, there has been interest from software developers in improving current image processing capabilities of flexible polymers. This will aid both academia and industry.

The research also led to the creation of 30 new monoclonal antibodies with differing specificities against AAT conformers. These have the potential to elucidate new information about disease, native and artificially-made protein. One antibody, 1D9, was characterised in-depth and found to be unique in its specificity for the monomeric form of AAT. This has direct medical relevance in diagnosis, as patients with AATD have low levels of monomer in their bloodstream. Currently, AATD is underdiagnosed and 1D9 offers the possibility of commercialisation, to compliment or supersede the existing tools for diagnosis. To progress this, together with therapeutic possibilities, pharmaceutical companies are heavily engaged with the laboratory.

This thesis contributes to the understanding of AATD and related diseases by using EM to characterise the structure of the AAT polymer, as well as developing diagnostic tools and therapeutic approaches. The work also provides a strong foundation for colleagues to continue in the future.

Table of contents

Table of figures

Abbreviations

2D	Two dimensional
3D	Three dimensional
AAT	$Alpha_1$-antitrypsin
AATD	$Alpha_1$-antitrypsin deficiency
ADC	Antibody drug conjugate
AFM	Atomic Force Microscopy
CCD	Charge-coupled device
CDR	Complimentary determining regions
CG	Covent Garden
CMS	Culture media supernatant
CPS1	Carbamoyl phosphate synthetase 1
CTF	Contrast Transfer Function
DED	Direct electron detector
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
ER	Endoplasmic reticulum
ERAD	ER-associated protein degradation
FEG	Field emission gun
FENIB	Familial Encephalopathy with Neuroserpin Inclusion Bodies
Fab	Fragment antigen binding
FBS	Foetal bovine serum
Fc	Fragment constant
FR	Framework region
FRET	Fluorescence resonance energy transfer
FSC	Fourier shell correlation
FT	Fourier transfer
Gu-HCl	Guanidine Hydrochloric acid
HDAC	Histone deacetylase
hIPSC	human Induced pluripotent stem cells
HNE	Human neutrophil elastase
HRP	Horseradish peroxidise
IB	Inclusion body
Ig	Immunoglobulin
mAb	Monoclonal antibody
ML	Maximum likelihood
MRA	Multi-reference alignment
MS	Mass spectrometry
MSA	Multivariate statistical analysis
MWCO	Molecular weight cut-off

NMR	Nuclear Magnetic Resonance
NS	Negative stain
Ox-AAT	Oxidised antitrypsin
PAI-1	Plasminogen activator inhibitor-1
PAGE	Polyacrylamide gel electrophoresis
PAS	Periodic acid-Schiff
PBA	4-phenylbutyric acid
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PSF	Point spread function
RCL	Reactive centre loop
RMSD	Root mean squared deviation
RT	Reverse transcription
RU	Response units
SAXS	Small angle X-ray scattering
SDS	Sodium dodecyl sulfate
SI	Stoichiometry of inhibition
SPR	Surface plasma resonance
SNR	Signal to noise ratio
TEM	Transmission electron microscopy
UA	Uranium acetate
UPR	Unfolded protein response

1.0 Introduction to Alpha$_1$-Antitrypsin

1.1 The serpin superfamily

Serpins are the largest and most broadly distributed family of protease inhibitors. Homologues have been identified in all domains of life- animals, plants, bacteria, archaea and viruses with a total of more than 1,500 members[1]. The majority of serpins are serine protease inhibitors, a role after which the family is named, however some have activity against cysteine proteases. The inhibitory serpins have important roles in the regulation of proteolytic cascades such as those in the inflammatory (alpha$_1$-antitrypsin, AAT), complement (C1-inhibitor), coagulation (antithrombin) and fibrinolytic (plasminogen activator inhibitor-1, PAI-1) pathways[2]. Other members are non-inhibitory with roles ranging from blood pressure regulation (angiotensinogen) to chromatin condensation (MENT)[2,3]. Some of the serpins reffered to in this thesis are listed in Table 1.1.

Serpin	Location	Activity	Function
Alpha$_1$-antitrypsin	Extracellular	Inhibitory	Inhibition of Neutrophil elastase (inflammation)
Angiotensinogen	Extracellular	Non-inhibitory	N-terminal cleavage releases the hormone angiotensin 1 (blood pressure regulation)
Antichymotrypsin	Extracellular	Inhibitory	Inhibition of Cathepsin G (inflammation)
Antithrombin	Extracellular	Inhibitory	Inhibition of Thrombin and Factor Xa (blood clotting)
C1-inhibitor	Extracellular	Inhibitory	Inhibition of C1 esterase (complement cascade)
Corticosteroid-binding globulin	Extracellular	Non-inhibitory	Transport of hormones, particularly cortisol
Heat shock protein 47	Intracellular	Non-inhibitory	Chaperone for collagens
MENT	Intracellular	Both	Chromatin condensation and inhibition of cathepsin L (protein degradation)
Plasminogen activator inhibitor-1	Extracellular	Inhibitory	Inhibition of thrombin, uPA, tPA and plasmin (fibrinolysis)
Plasminogen activator inhibitor-2	Intracellular	Inhibitory	Inhibition of uPA (fibrinolysis)
Protease nexin 1	Extracellular	Inhibitory	Inhibition of uPA and tPA (fibrinolysis)
Neuroserpin	Extracellular	Inhibitory	Inhibition of tPA, uPA and plasmin (fibrinolysis)

Table 1.1. A selection of human serpins with their location and function. *Table adapted from [4] and supplemented with [5,6].*

1.2 The serpin fold

Serpins have a conserved tertiary fold despite a relatively low sequence identity[7] (Figure 1.1). Their structure consists in most cases of 9 α-helices (sequentially named A to I), 3 β-sheets (named A to C) and a reactive centre loop (RCL). The RCL is typically 21 amino acids long, labelled P17-P4' and in the native state is solvent-exposed and mobile.

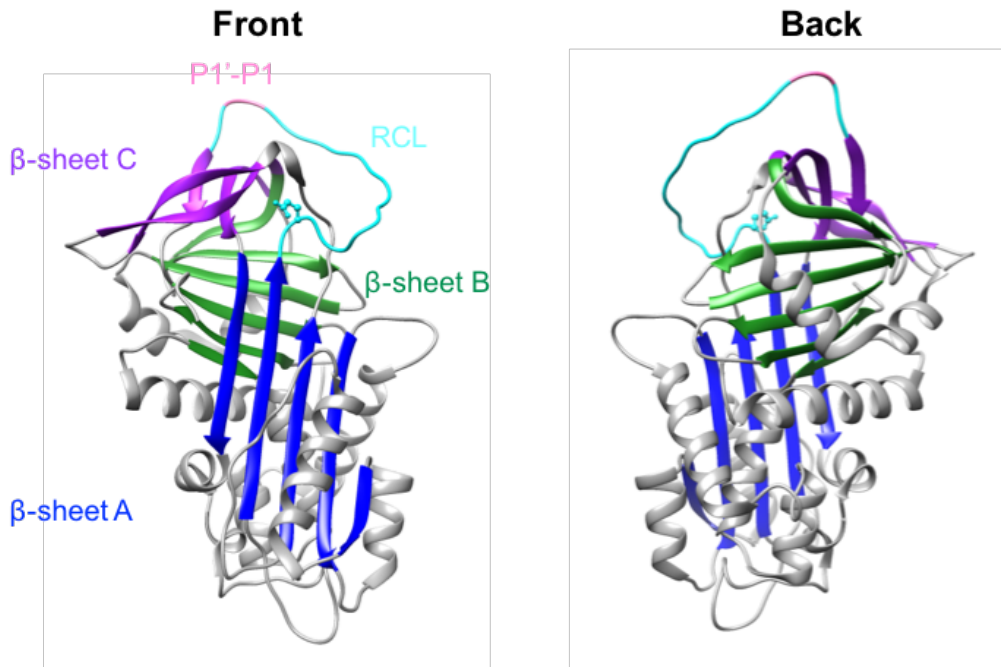


Figure 1.1. The serpin fold. *Front (left) and back (right) views of AAT (PDB ID:1QLP) shown in ribbon representation[8]. The 3 β-sheets are highlighted with β-sheet A in blue, β-sheet B in green and β-sheet C in purple. The RCL is shown in turquoise with the site of the common pathological 'Z' mutation shown in ball and stick representation and the P1'-P1 protease cleavage bond in pink. Figure made using Chimera[9].*

Serpins are an exception to Anfinsen's principle[10]: they fold into a kinetically trapped, metastable or 'stressed' state that can readily change conformation, rather than the conformation with the lowest free energy. This is important for serpin function but also renders the serpins susceptible to mutations that lead to misfolding and polymerisation. For inhibitory serpins, the metastability is predominantly resolved upon insertion of the RCL as a 6th strand in the central mixed parallel/anti-parallel five-stranded β-sheet A[11]. This makes β-sheet A completely antiparallel, forming the 'relaxed' serpin conformation. RCL insertion releases around 50-60 kcal/mol and the thermal stability of the serpin increases from 50-60℃ to greater than 120℃[12].

1.3 The serpin mechanism of protease inhibition

The RCL of inhibitory serpins is highly flexible and its conformation has important biological and functional implications. Its movement is particularly central to the 'suicide-substrate' and irreversible mechanism of protease inhibition (Figure 1.2)[13]. During inhibition, the RCL acts as a pseudosubstrate by presenting the P1-P1' residues to the target protease in an optimal conformation[14]. In the case of AAT, the P1-P1' residues are Met358-Ser359[3]. Cleavage by the protease at the P1'-P1 bond leads to the formation of an acyl-intermediate, with an ester bond linking the protease to the serpin P1 residue[15]. Release of the C-terminal fragment permits the RCL to fully insert as a 6th strand (strand 4A) in β-sheet A, flipping the protease 70Å to the opposite pole of the serpin and adopting the relaxed state[13,16,17]. In the process, the protease active site serine is pulled out of the catalytic triad conformation, leading to distortion of the active site and a 37% loss of structure in the protease through compression against the serpin[13,18,19]. The protease mechanism therefore cannot proceed to deacylation of the complex. This process is thermodynamically driven, with the resultant serpin in the lowest free energy conformation and energy transferred to denaturation of the protease and converting the protease back into a zymogen-like state[13]. The level of denaturation observed may vary between serpin-protease pair as a crystal structure of porcine pancreatic elastase in complex with AAT showed no disorder to the protease[1,20]. The serpin-protease complex is then cleared quickly from the circulation by binding to lipoprotein receptors[21].

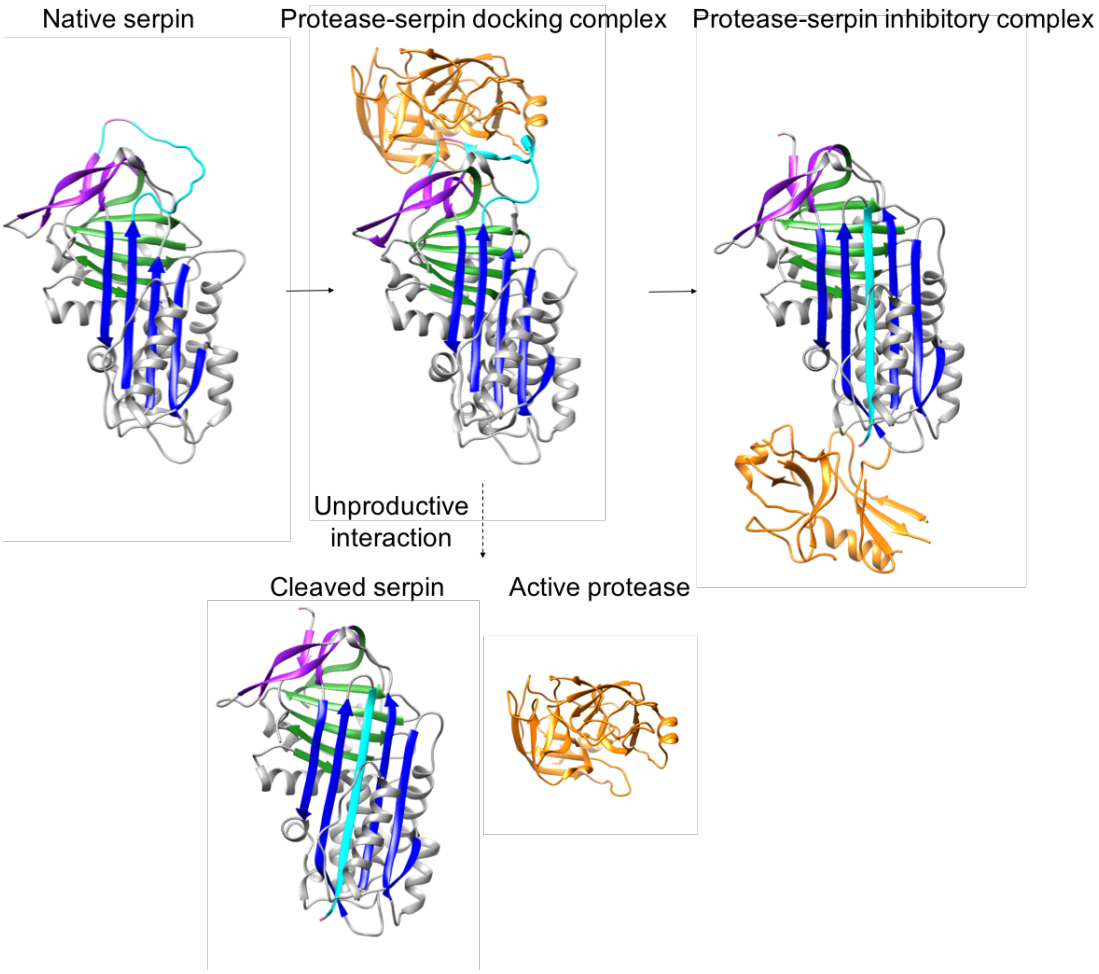


Figure 1.2. The serpin mechanism of protease inhibition. *The native conformation of the serpin (PDB ID: 1QLP) interacts non-covalently with the protease (orange) to form the docking complex (PDB ID:1OPH)[8,22]. The serpin RCL (turquoise) is cleaved by the protease and inserts into β-sheet A (blue). In a productive interaction, RCL-insertion translocates the protease to the opposite pole of the serpin, denatures the protease active site and forms a stable irreversible complex (PDB ID: 1EZX)[13]. In an unproductive interaction, the serpin is left inactive in a cleaved conformation and the active protease is released. Figure made using Chimera[9].*

1.4 Other hyperstable serpin conformations

Some serpin-protease pairs will result in a proportion of unproductive interactions. This is where the protease cleaves the serpin but is not covalently bound and therefore is not irreversibly inhibited. The serpin-protease complex dissociates; the protease is released to cleave other proteins but the serpin is left inactive (Figure 1.2). The released serpin has a cleaved RCL, which similarly is fully self-inserted to complete β-sheet A (Figure 1.3, left)[14]. The cleaved serpin is thermally hyperstable.

The RCL can also be incorporated into β-sheet A in the absence of cleavage, forming the latent conformation, which is associated with a loss of inhibitory activity (Figure 1.3, middle)[23]. The formation of the latent state requires the release of strand 1C. This locally destabilises β-sheet C and is proposed to reduce the global stability of latent compared to the cleaved conformation[3]. The latent conformation can be restored to the native conformation by denaturation and refolding. Some serpins adopt the inactive latent conformation *in vivo:* latent antichymotrypsin and neuroserpin have been detected in individuals[24], however the latency transition has been best characterised in plasminogen activator inhibitor-1 (PAI-1). PAI-1 undergoes this process by default, with a half-life of 1 to 2 hours; vitronectin, a glycoprotein present in the blood, stabilises against the transition, increasing the half-life by between 2-16 fold depending on buffer conditions[25]. AAT can be artificially induced to form the latent conformation by heating in 0.7M citrate but this form has not been detected *in vivo* except in individuals who have received augmentation therapy, due to presence in the administered AAT[26]. The lack of latent AAT *in vivo* may be due to the relatively short RCL that does not permit complete formation of strand 4A[27].

RCL insertion into β-sheet A can also be incomplete or partial, as captured in a crystal structure of a mutant antichymotrypsin atypically stabilised by the loop C-terminal to helix F (Figure 1.3, right)[28]. In this publication, the structure was proposed to be the polymerisation intermediate, having a stability between native and latent conformations. Partial insertion has also been observed with non-mutant antichymotrypsin and antithrombin[29-31]. Notably both serpins have an RCL that is longer than that of AAT.

Lastly, the RCL β-sheet A interaction is proposed to be the basis for the aberrant intermolecular linkage present in serpin polymers that form as a result of mutations[32]. These polymers are highly stable and can have pathological consequences *in vivo*, and will be discussed in detail below.

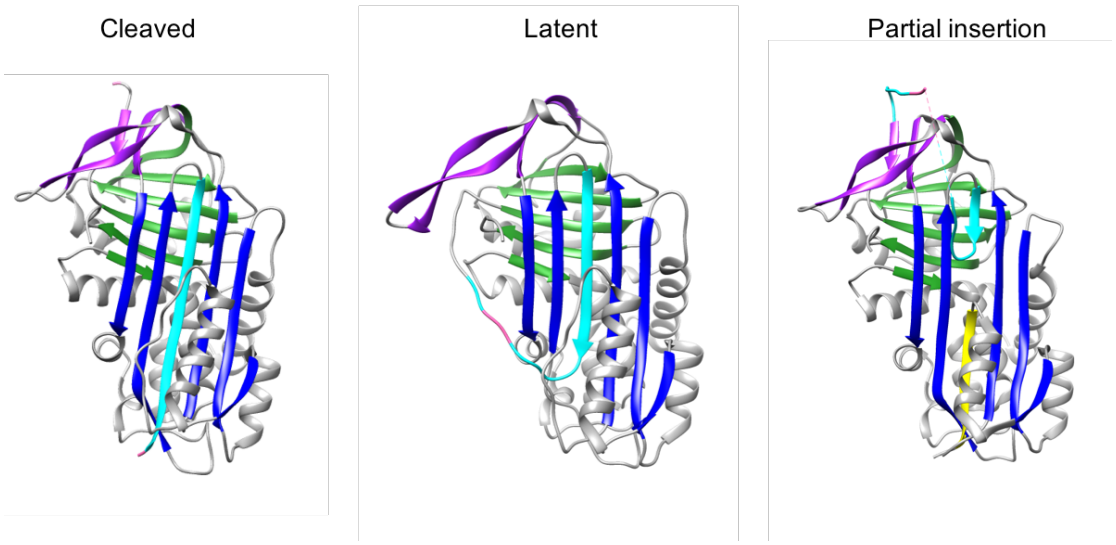


Figure 1.3. Hyperstable serpin conformations. *The cleaved conformation (PDB ID: 1EZX) is formed by cleavage of the P1'-P1 bond and insertion of the RCL into β-sheet A but with dissociation of the protease*[13]*. The latent conformation (PDB ID: 1IZ2) is formed by insertion of the RCL in the absence of cleavage*[27]*. The RCL can also be partially inserted into the top of β-sheet A (PDB ID: 1QMN), with the bottom stabilised by unwinding of helix F into a β-strand (yellow)*[28]*. Figure made using Chimera*[9]*.*

1.5 Alpha$_1$-antitrypsin (AAT)

Alpha$_1$-antitrypsin (AAT) is a 52 kDa glycoprotein consisting of 394 amino acids. It is the most abundant protease inhibitor in the plasma at a concentration of 1.5-3.5g/l in a healthy individual[3]. AAT is an acute-phase protein involved in inflammation so levels can transiently increase 4-fold above this range[3]. It has a half-life of 3-5 days in serum, strongly dictated by the presence of its glycan chains: removal of one glycan reduces the half-life significantly[33].

Over 80% of AAT is synthesised in hepatocytes[3]. Normal physiological production of AAT involves co-translational translocation into the endoplasmic reticulum (ER) and glycosylation of three asparagine residues with branched, 14-unit oligosaccharides[34]. These appendages direct AAT into a specialised glycoprotein folding pathway aided by chaperones, where the glycans are trimmed and altered by various glucosidases[35]. On maturation, AAT is transported to the Golgi complex where further glycan processing occurs[35]. It is then secreted into the blood for transportation to the lungs. A small amount of local secretion of AAT also occurs from lung epithelial cells, neutrophils and alveolar macrophages[36-38]. Within the lungs, the main target of AAT is neutrophil elastase secreted by neutrophils at sites of inflammation.

It also inhibits proteinase-3[39]. AAT inhibition of these enzymes serves to protect the alveoli from unregulated elastin breakdown.

1.6 Alpha$_1$-antitrypsin deficiency (AATD) and the serpinopathies

Mutations within the AAT gene can alter the normal physiological production of AAT and result in disease. AAT deficiency (AATD) is one of many conformational diseases that result from the polymerisation of mutants of members of the serpin superfamily. Collectively, these are known as the serpinopathies[40]. Each serpinopathy is associated with a gain-of-function phenotype caused by the polymerisation and retention of the serpin at the site of synthesis, or a loss-of-function phenotype caused by the resultant deficiency, or both. For example, polymerisation of neuroserpin in the cerebral cortex results in an autosomal dominant dementia known as Familial Encephalopathy with Neuroserpin Inclusion Bodies (FENIB), whilst the deficiency caused by polymerisation of antithrombin results in thrombosis[41]. AATD displays both gain-of and loss-of-function effects, with the accumulation of polymers in hepatocytes causing liver disease and the deficiency in the plasma causing emphysema[42].

AATD is an autosomal co-dominant disorder and is associated with many disease-causing mutations in the SERPINA1 gene (encoding AAT). Over 100 naturally-occurring allelic variants of AAT have been identified and over 30 of these affect the *in vivo* function or secretion of AAT (Figure 1.6)[43]. AAT variants have in the past been classified according to their migration by isoelectric focusing (Figure 1.4), where proteins are separated by charge and stop migrating at the pH when their net charge is zero. However, modern convention is to name variants according to the geographical region or the hospital of their discovery. The wild-type AAT protein is referred to as M-AAT. The most common mutant variants are the 'Z mutation' or 'Z-AAT', with a Glu342Lys single point mutation and S-AAT, with the Glu264Val mutation.

Isoelectric focusing of plasma AAT has demonstrated that multiple glycoforms of AAT exist (Figure 1.4). Nine different glycoforms of M-AAT (classified as subtypes M0-M8) and six of Z-AAT have been identified[44]. The forms vary in the number of sialic acid residues on the glycan chains, as a consequence of different glycan branching patterns. Glycoanalysis comparing the isoforms of M-AAT and Z-AAT has

shown that there was no significant difference in the branching structure of the glycans, however there were more glycans on Z-AAT with a core and/or outer-branch fucose residue[45]. The exact effects of this is unknown but it has been hypothesised to alter AAT's anti-inflammatory properties.

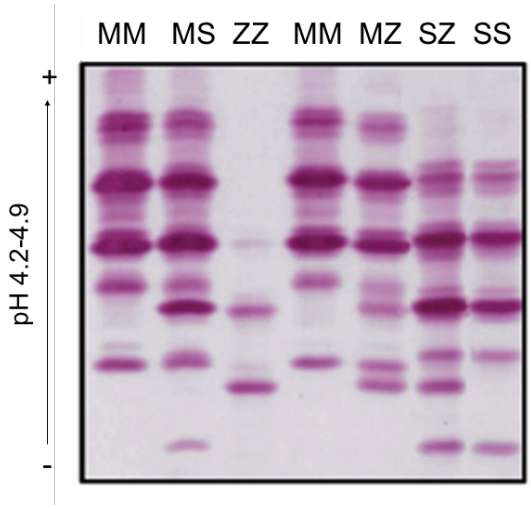


Figure 1.4. An example isoelectric focusing gel used for the identification of different AAT genotypes. *The multiple bands per AAT variant represent different glycoforms and N-terminal truncations of AAT. Figure adapted from[46].*

The Z mutation is situated at the base of the RCL and head of s5A where it anchors s5A to s6A through a salt bridge with Lys290[47]. The Z mutation is thought to promote polymerisation by further destabilising the native metastable conformation and promoting formation of the β-sheet A-opened intermediate[48]. Glu342 has been shown to be critical for proper folding, and a positive charge at this position, such as the Z mutation, reduces folding efficiency. However, the crystal structure of glycosylated Z (Glu342Lys) AAT expressed in Drosophila resembles the wildtype structure with minimal effects of Lys342 on packing around the hinge[49]. Similarly, low resolution AFM data found no major topographical differences between M and Z monomers[50]. A Glu342Ala AAT variant with a disulphide bond between s5A and s6A behaved similar to wildtype AAT in peptide binding studies, thermal stability and basal fluorescence, whereas the reduced form was similar to the Z form. Furthermore, residues on top of s5A on a Glu342Ala background, but not wildtype, were accessible for modification in a pegylation assay. These data indicate that the Z mutation results in an unstable β-sheet A in which the top half of s5A is labile[51]. A study that employed hydrogen/deuterium exchange combined with mass

spectrometry (MS), to probe the dynamic differences between M-AAT and Z-AAT, similarly concluded that the Z mutation perturbs the region around the mutation and even more distant regions such as B, D and I helices and β-sheet A[52]. Therefore, an aberrant conformation of Z-AAT likely arises from the loss of stabilising interactions of Glu342[48,49]. Other mutations are proposed to affect the sliding movement involved in the opening of β-sheet A or affect the movement of the RCL[53-55]. Equivalent AATD-causing mutations have been found in other disease-causing serpins[41].

More than 95% of individuals with severe AATD are homozygous for Z-AAT. The Z mutation is most common in Northern European populations occurring at a frequency of 1 in 40, and therefore affecting 1 in 1700[56]. The S allele occurs in 1 in 5 Southern Europeans, peaking in the Iberian Peninsula[43,56]. Individuals that are homozygous for the Z mutation retain 85-90% of synthesised AAT. The majority (~70%) of this protein is degraded by the ER-associated protein degradation (ERAD) pathway, which involves retrotranslocation of misfolded proteins from the ER lumen to the cytoplasm for degradation by the ubiquitin-proteasome pathway[57]. The remainder of the protein self-associates into polymers that are incorporated into insoluble, periodic acid-Schiff (PAS) positive inclusions (Figure 1.5)[32,57]. Consequently, there is a circulating deficiency of AAT monomer, and reduced antiprotease reaching the lungs. The S allele is less severe, with S homozygotes having AAT plasma levels of about 60% of that of M homozygotes[55].

Patients with AATD can present clinical symptoms involving both the liver and the lungs. AAT polymerisation within the hepatocyte ER causes cellular damage, hepatocyte proliferation and sensitisation to ER stress[59,60]. This predisposes individuals to neonatal hepatitis, liver cirrhosis and an increased risk of hepatocellular carcinoma. The deficiency of circulatory AAT predisposes to early-onset emphysema. This is due the over-activity of neutrophil elastase which breaks down pulmonary connective tissue and alveoli. Furthermore, the little aberrant AAT that reaches the lungs has been proposed to polymerise and co-localise with neutrophils, triggering an inflammatory response which causes additional tissue damage[61]. There is also some evidence that AAT activity is further compromised in individuals who smoke due to the oxidation of the P1 methionine residue to a sulfoxide[62,63].

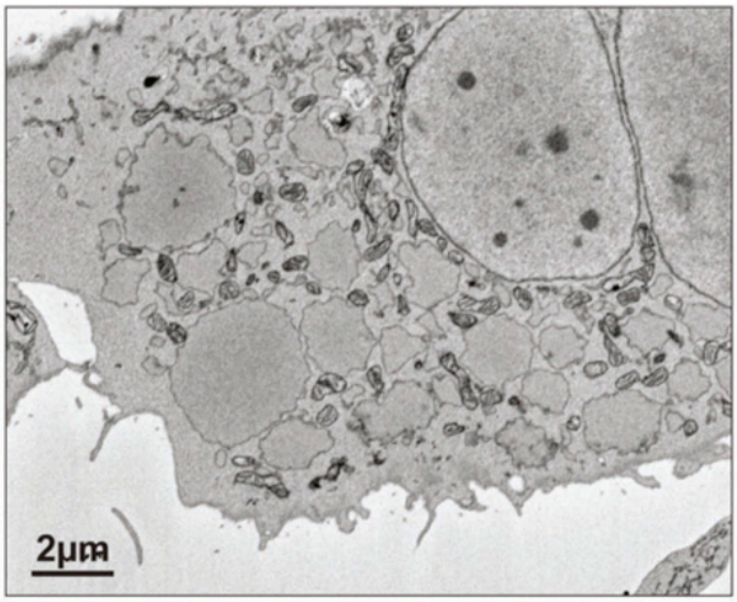
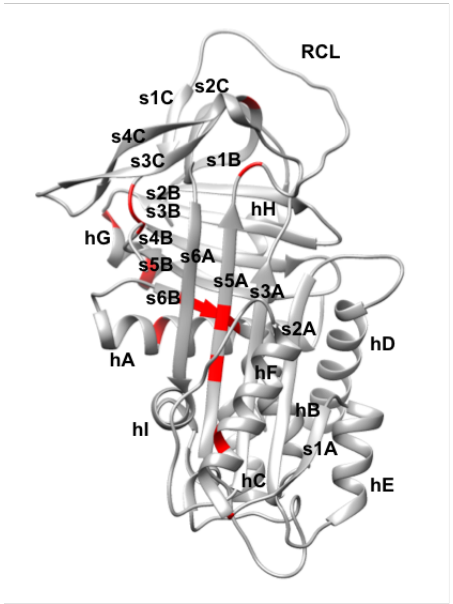


Figure 1.5. Inclusion bodies in CHO cells transfected with Z-AAT. *A single serial block-face scanning electron micrograph of a CHO cell with multiple rounded inclusion bodies is shown. Figure reproduced from* [58].



AAT Variant	Mutation	Structural location	Geographical predominance	Polymerisation? Liver disease?	Deficiency? Emphysema?
Baghdad	Ala336Pro	Middle s5A- "shutter" region	Single case of Iraqi origin	Yes, unknown	Yes, unknown
$E_{Taurisano}$	Lys368Glu	Loop between s4B and s1C- "gate" region	Cases found in Southern Europe	Yes in cell model, unknown	Yes in ME_T, yes
I	Arg39Cys	Helix A		Yes, yes in IZ	Yes, yes in IZ
Kings	His334Asp	Middle s5A- "shutter" region	Single case of European origin	Severe in KZ, yes	Severe in KZ, presumed
$M_{Heerlan}$	Pro369Leu	Loop between s4B and s1C- "gate" region		Yes, unknown	Severe in MhZ, yes
M_{Malton}	ΔPhe51	s6B- "shutter" region	Sardinia	Yes, yes	Yes, yes
$M_{Nichinan}$	ΔPhe52 + Gly148Arg	s6B/helix B- "shutter" region + Loop between s1A and helix F	Sardinia	Yes, unknown	Yes, unknown
M_{Pisa}	Lys259Ile	Helix G- "gate" region	Cases found in Southern Europe	Yes in cell model, unknown	Yes in M_PZ, yes
$P_{Brescia}$	Gly225Arg	Loop between s1B and s3C	Cases found in Southern Europe	Yes in cell model, abnormal in P_BX	Yes in P_BX, presumed
Queen's	Lys154Asn	Helix F		Polymerisation *in vitro*, unknown	Mild, unknown
S	Glu264Val	Helix G-"gate" region	Iberia, Europe	Mild, yes in SZ	Mild, yes in SZ
Siiyama	Ser53Phe	Helix B- "shutter" region	Japan	Yes, unknown	Yes, yes
Wurzburg	Pro369Ser	Loop between s4B and s1C- "gate" region	Case found in Europe	Yes in cell model, unknown	Yes, unknown
$Y_{Orzinuovi}$	Pro391His	End of s5B	Cases found in Southern Europe	Yes in cell model, abnormal in MY_O	Mild in MY_O, presumed
Z	Glu342Lys	Top of s5A, base of RCL- "hinge" region	Northern Europe	Severe, yes	Severe, yes

Figure 1.6. Polymerogenic mutations of AAT. *A) The structure of AAT with mutations coloured in red. B) A summary of the properties of the polymerogenic AAT variants. Table reproduced from* [42] *and supplemented with* [64,65].

1.7 Biochemical characterisation of polymerisation of AAT *in vitro*

As polymerisation of AAT underlies the disease, much work has focused on characterising the pathway to polymerisation and the structure of terminal polymers. The majority of experiments have studied the artificial generation of polymers from purified monomeric protein. Polymerisation can be induced under multiple different mild-denaturing conditions: by elevated temperature, incubation with denaturants such as Guanidine HCl (Gu-HCl) and urea, acidic pH and by cleavage of the RCL in non-standard positions[47,66-69]. The topography and periodicity of the polymers have been found to vary according to the method of polymerisation[50]. They also have different appearances on a non-denaturing gel[69,70].

Heat-induced polymers are formed by heating AAT between 41-65°C[32,71-74]. A multitude of spectroscopic techniques have suggested that heat-induced polymerisation is biphasic (Figure 1.7)[71,72]. Fluorescence studies using bis-ANS showed a fast initial increase in fluorescence followed by a slower reduction in fluorescence signal[71]. Similar findings were observed with intrinsic tryptophan fluorescence, fluorescence polarisation, circular dichroism and extrinsic fluorescence using a rhodamine probe and the ANSA dye[72]. Calculation of rates showed that the first step was unimolecular (concentration-independent) and 10-fold faster than the bimolecular polymerisation step under the experimental conditions used[72].

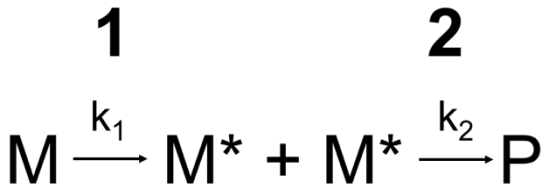


Figure 1.7. Kinetic model for the polymerisation of AAT. *Step 1- An unstable intermediate (M*) is formed from monomer (M) by rate k_1. Step 2- M* molecules associate to form polymers (P) by rate k_2.*

The first step has been attributed to formation of an unstable but activated intermediate (M*)[72]. This may involve the opening of β-sheet A, potentially involving partial RCL insertion and s1C movement. Polymerisation did not occur when s1C was bound to s2C using a disulphide bond, in keeping with the conformation

observed for the antichymotrypsin intermediate[28,71,73,74]. A monomeric-sized species, proposed to be M*, has been characterised by MS[69]. It was recognised by the polymer-specific monoclonal antibody 2C1, which suggests it shares a conformational epitope present in monomeric subunits within the polymer, but absent in the native monomer. M* had a larger collision cross section than the monomer in ion mobility MS indicating the M* has an enlarged surface area compared to the monomer. The second phase has been attributed to insertion of the RCL during polymerisation. This kinetic mechanism is supported by the lag in polymer formation as the concentration of M* rises observed by non-denaturing gels and rhodamine fluorescence[72]. Dimers have been proposed to propagate polymerisation by having one exposed but partially inserted RCL as a β-strand donor and a readily opened β-sheet as an acceptor[75].

Heat polymers are recognised by the monoclonal antibody 2C1, which also recognises polymers found extracted from hepatocytes of individuals with AATD[76]. They therefore share an epitope and are thought to be most physiologically relevant; neither polymers formed by incubation in denaturant or acid are recognised by 2C1[69]. Denaturant polymers can be formed by incubation with 1-3M Gu-HCl or 4M urea[32,47,69]. Gu-HCl is thought to induce unfolding to form an intermediate that has local unfolding in the s5A-hI-s6A and F helix areas and incubation results in a high proportion of dimers rather than high-molecular weight polymers[68,77]. Acidic polymers have been formed by incubation with 0.1M sodium acetate, pH 4-4.5. They are proposed to form via an intermediate with a rearrangement at the top of β-sheet A and strand 1C and overall a 20% loss in secondary structure[66]. This is similar to the intermediate formed by heating, however polymerisation at low pH follows a three-stage mechanism in contrast to the two-stage mechanism of heat-induced polymerisation[66,71]. Incubation at low pH formed the self-terminating antithrombin structure that underlies the β-hairpin model of polymerisation[78]. Lastly, polymers can be formed by cleavage of the RCL at P5-P4, P7-P6, P9-P8 or P10-P9[67]. Residues N-terminal to the cleavage site insert into the top of β-sheet A, opening up the bottom of the sheet for insertion of residues from another cleaved monomer[79,80]. Cleaved polymers however are unlikely to be physiological as *in vivo* polymers have an intact RCL.

1.8 The models of polymerisation

Polymers formed *in vivo* arise from a domain swap between consecutive AAT subunits and are hyperstable. This is consistent with an RCL-inserted 6-stranded β-sheet A, and notably the RCL is protected from proteolysis on polymerisation[32,67,81]. Subunits are also non-covalently linked as polymers can be readily dissociated by 1% w/v SDS[81].

The pathological mutations of AAT that lead to polymerisation occur in structurally distinct regions of AAT. The polymer-specific monoclonal antibody called 2C1 recognises polymers resulting from a variety of mutations[76,82]. This indicates the polymers formed by different mutations *in vivo* share an epitope, which increases the likelihood that they are structurally similar.

The structure of the *in vivo* AAT polymer is unknown. Previous studies using electron microscopy (EM) have shown that polymers are flexible and predominately linear, with beads-on-a-string morphology[32,81]. Polymer circlets are occasionally observed and their presence indicates that donor and acceptor ends are complementary, allowing self-terminating polymers to form. The existence of self-terminating and ordered polymers may explain why polymer accumulation does not activate the unfolded protein response (UPR)[60]. Early fluorescence studies using bis-ANS suggested that the polymer structure has newly exposed hydrophobic surfaces relative to the monomer and AAT must undergo a significant structural change upon polymerisation[71]. Several models of polymerisation have been proposed that differ in the intermolecular linkage between monomers (Figure 1.8):

(A) The loop-β sheet A model[32]. This model proposes that the RCL of one monomer inserts into the β-sheet A of another.

(B) The loop-β sheet C model[83,84]. This model proposes that the inter-subunit interaction occurs between the RCL of one molecule and the β-sheet C of the next.

(C) The strand 7A model[5,85]. Here, the RCL forms an extra lateral strand of β-sheet A.

(D) The β-hairpin model[78]. In this model, the intermolecular linkage is a β-hairpin, formed by the RCL and neighbouring strand 5 of β-sheet A.

(E) The C-terminal model[86]. The intermolecular linkage consists of three β-strands, formed by strand 1 of β-sheet C and strands 4 and 5 of β-sheet B.

All models except (B) agree that the β-sheet A of AAT transitions from being 5-stranded to 6-stranded. This explains the hyperstability of polymers and the protection of the RCL from proteolysis[3,67]. However, the models disagree on the involvement of strands which engage in the intermolecular linkage. Each model is based on a combination of structural and biochemical/biophysical data, generated from polymers made artificially by different *in vitro* methods or expressed in cells without glycan chains. Therefore, which, if any, model is applicable to the pathological context is currently unknown and this question will be addressed in this thesis. In fact, AFM studies comparing *in vitro* formed polymers (made by heating, acid or Gu-HCl incubation) to Z polymers isolated from the liver of a transgenic mouse, showed that long smooth linear polymers were a prominent and unique component of the sonicated insoluble fractions of liver extracts[50]. The loop-β sheet A, β-hairpin and C-terminal models have had the most backing for the polymer structure, although it is possible that multiple forms of polymers may be present *in vivo*. Accordingly, oligomers of topographies indistinguishable from all *in vitro* polymers were present in the *ex vivo* liver polymer sample, as well as the long fibres unique to the mouse liver polymers[50].

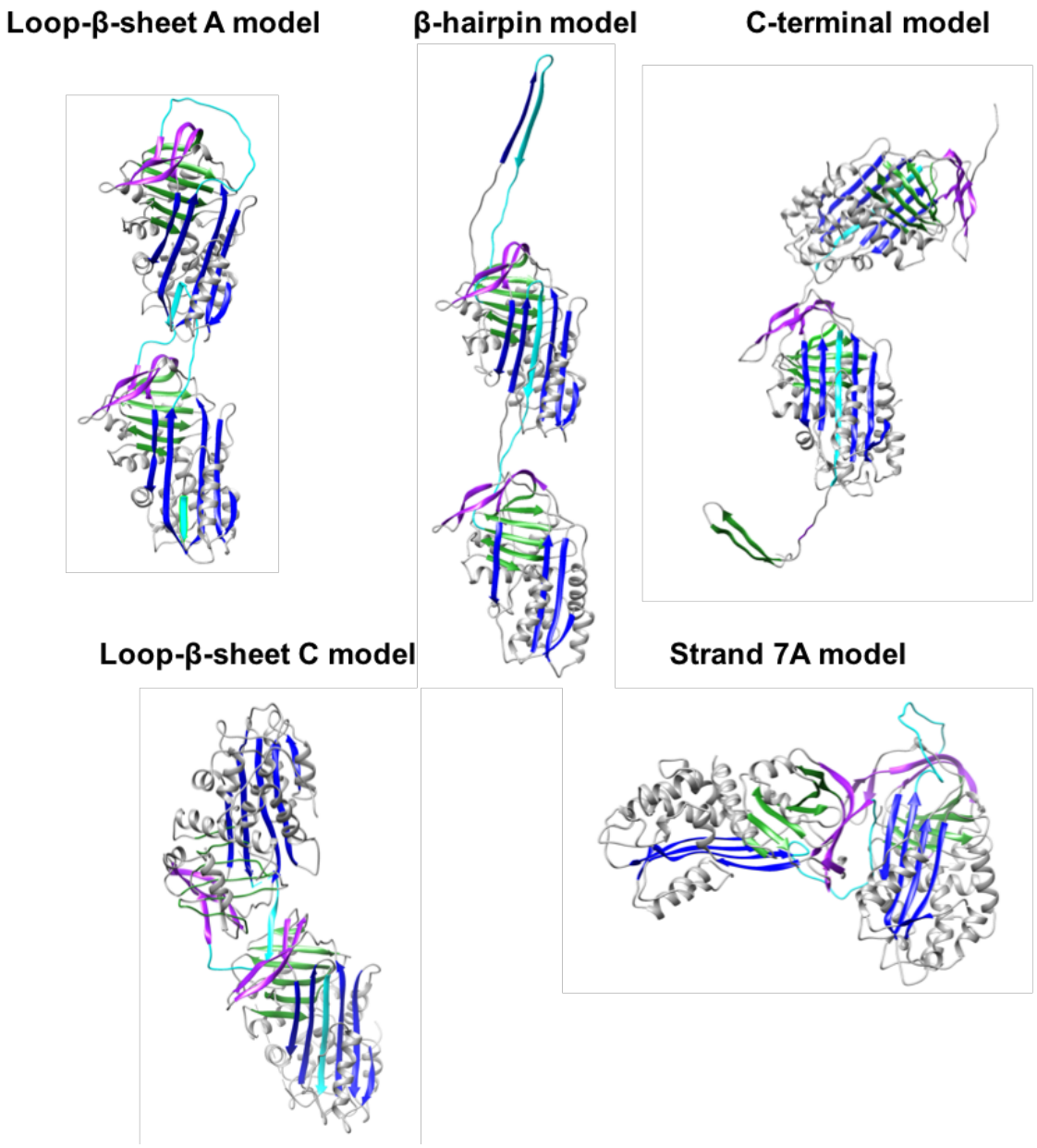


Figure 1.8. The proposed models of AAT polymerisation. *The loop β-sheet A model was made by superimposing native AAT (PDB ID:1QLP[8]) on the AAT:TTAI ternary complex (PDB ID: 4PYW[87]). Phenix was used to simulate the P3-P8 insertion. The loop β-sheet C model is based on PDB ID: 1ANT[83]. The strand 7A model is based on PDB ID: 1B3K[85]. The β-hairpin model is based on PDB ID: 2ZNH[78]. The C-terminal model is based on PDB ID: 1D5S[86]. For all models, β-sheet A is in magenta, RCL in turquoise, β-sheet B in orange and β-sheet C in green. Models generated by Dr. J. Irving.*

1.8.1 Evidence for the polymerisation models

A) The loop-β sheet A model

The loop-β sheet A model was the first proposed model of polymerisation and describes how the RCL of one molecule inserts into the β-sheet A of another[32]. The basis for this was the ability of RCL peptides to block polymerisation, which was later supported by crystal structures with peptides in the s4A position[32,88]. The number and specific residues of the RCL thought to constitute the intermolecular linkage has changed over the years. Initially, P12-P5 was thought to be the intermolecular linker, however the observation that P14-P9 peptides increase, rather than inhibit polymerisation, meant the model was revised to P8-P3 intermolecular insertion[53,73]. This requires s1C partially peeling away to increase mobility of the RCL and would leave an unfavourable gap at the top of β-sheet A. The gap is proposed to be satisfied by self-insertion of the RCL[73], based on the crystal structure of a polymerogenic antichymotrypsin mutant (a serpin related to AAT but with a longer RCL), where partial self-insertion (P14-P12) of the RCL and opening of β-sheet A was observed[28]. The structure showed strand s1C had been distorted and the lower part of the sheet was stabilised by the last turn of the F-helix and a loop linking it to s3A. This conformation was proposed to be the intermediate conformation (M*) on the polymerisation pathway, however this has been disputed by the observation that it represents a terminally misfolded form of antichymotrypsin[29]. P8-P3 intermolecular insertion is also supported by cleavage of AAT at P9-P8 or P10-P9 resulting in spontaneous polymerisation[67], although this cleavage would alleviate any steric hindrance imposed by an intact RCL.

The loop-β sheet A mechanism has been tested by RCL mutations. Mutation of hydrophobic residues P8-P6 to Asp was used to investigate whether these residues make up the majority of the intermolecular linkage, or are just a small part of a larger domain swap[89]. These positions are oriented towards the protein core upon RCL insertion and would be expected to render canonical incorporation of the RCL unfavourable. The AAT variant did polymerise when induced with Gu-HCl, with heat and in transiently transfected COS-7 cells expressing AAT with a Z background. Further mutations at P10 and P9 were still tolerated; the conclusion was that the intermolecular polymeric linkage is not comprised of a partially-inserted RCL and instead a larger domain swap must occur. Of note, the variant was unable to

polymerise when incubated with small polymerisation-inducing peptides or when the RCL was cleaved in a region that is known to create 'cleaved polymers'[80]. Another study similarly mutated residues along the RCL to probe the polymerisation pathway[90]. Circular dichroism, bis-ANS studies and a FRET-based assay suggested that P14P12 self-inserts early during the pathway, whereas P6P4 interact with β-sheet A (either inter- or intra-molecularly) later in polymerisation. Furthermore, only mutations at P14P12 and P4, but not at P10, P8 or P6, resulted in a decrease in polymer levels in a transiently transfected COS-7 cell model. The authors argued that in order for these residues to be unfavourable for polymerisation, the AAT molecule would need to be substantially folded and hence that polymerisation proceeds from a 'near-native' conformation. Similarly, a separate NMR study that characterised the intermediate ensemble formed by the Queen's (K154N) mutation at physiological conditions found a high correlation between M-AAT and Queen's AAT resonances, arguing against an unfolded intermediate[82]. While these studies do not rule out the linkages proposed by the β-hairpin and C-terminal models, it contradicts the nature of the intermediates hypothesised by those studies.

Unlike the other models, there is no crystal structure of an intact loop-β sheet A polymer. However, the loop-β sheet A model has been evaluated for physical plausibility by computationally creating pentameric molecular models of AAT[91]. Increased RCL insertion in *trans* resulted in predicted polymer compaction and rigidity, in contrast to the flexible polymers observed by EM. P11-P3 was suggested to be the insertion limit for maintaining reasonable stereochemistry and reducing molecular stress. RCL self-insertion on a P8-P3 template (where P8-P3 is the inter-molecular linkage and provided by the neighbouring subunit) was also modelled but self-insertion past P15 led to a highly compacted and strained polymer. Therefore, assuming that nearly complete β-sheet A occupancy is required for stability, self-insertion to form M* and polymers was concluded by the authors to be implausible and the rigidity of the polymer formed from P11-P3 insertion concluded to be unlikely[91]. However, the extent of RCL insertion required to achieve the stability seen in polymers has not yet been tested.

Different extents of a loop-β-sheet A intermolecular linkage have been observed in the crystal structures of cleaved polymers. An AAT polymer cleaved at the P5-P6 position of the RCL had the P17-P6 region self-inserted into β-sheet A and the P5-P1 region forming the β-strand linkage at the bottom of the β-sheet A of another

monomer[80]. A crystal structure with cleavage at P6-P7 showed self-insertion of P14-P7 and the intermolecular linkage formed by P6-P3 into the β-sheet A of another[79].

B) The loop-β sheet C model

This model was an alternative loop-sheet polymerisation mechanism originally proposed from the 3Å dimeric structure of antithrombin[83]. The structure showed one active molecule with an RCL self-inserted up to P14 and its P3-P7 residues replacing strand 1C of the other molecule, which was in the latent conformation. The authors postulated that polymers could form from a molecule with an opening β-sheet A that starts to incorporate its own RCL (i.e. on the transition to the latent state). This requires release of strand 1C and presents an incomplete β-sheet C to another molecule to allow polymerisation. Fluorescence studies have suggested that AAT can polymerise via the loop-β sheet A mechanism or the loop-β sheet C mechanism depending on the buffer conditions[92]. The latter is favoured by citrate buffer which promotes transition to the latent state, however complete formation of the latent conformation would block further polymerisation. This is proposed to explain the longer polymers formed by AAT compared to antithrombin, where antithrombin undergoes the transition to latency more readily.

A similar loop-β sheet C interaction was observed in the crystal lattice of an N-terminally truncated variant of a bacterial serpin called tengpin[84]. The RCL of tengpin was hyperinserted into its own β-sheet A, releasing s1C and presenting an exposed s2C for strand interactions with an exposed region of the RCL of a neighbouring molecule. In this structure, the strand interaction was antiparallel whereas in the antithrombin dimer the RCL of the active molecule linked in a parallel manner. These dimers dissociated readily to monomers by gel filtration in contrast to the stability expected of physiological polymers. The loop-β sheet C polymers are therefore thought to represent polymers formed by latent serpins e.g. S49P neuroserpin and PAI-1[24,84]. However, the truncation of the N-terminus in this serpin was found to favour the latent conformation and the formation of this conformation may be due to crystal packing.

C) The strand 7A model

A strand 7A interaction was first observed in the crystal structure of a quadruple variant of active PAI-1[85]. PAI-1 has a lower barrier to latency compared to other

serpins and spontaneously transitions to the latent conformation[23]. To counteract this in crystallisation trials, mutations were made in regions to stabilise the active form. In two out of the four molecules in the unit cell, the RCL was disordered and flexible. In the other two molecules, the RCL was constrained by crystal-packing contacts and added an extra β-strand (s7A) to β-sheet A. Similar to the loop-β sheet C polymers, the crystals could be dissolved to give active PAI-1, indicating the interaction was reversible.

The s7A loop-sheet interaction was also observed in the crystal lattice of a variant (ΔMloop) of MENT, a serpin that plays a role in chromatin condensation[5]. MENT can be crosslinked in oligomers in complex with DNA suggesting a functional role of polymerisation. A 12-mer RCL peptide was able to inhibit DNA condensing activity *in vitro*[5]. Polymerisation was probed by incubation at 44℃ for 16 hours and consistent with the s7A loop-sheet interaction, polymerisation was reversible and the oligomers retained inhibitory activity, whereas incubation at 55℃ resulted in the formation of typical inactive polymers. Therefore, the s7A loop-sheet model may represent a reversible physiological linkage rather than a pathogenic linkage.

D) The β-hairpin model

The β-hairpin model is based on the crystal structure of a self-terminating antithrombin dimer solved to 2.8Å[78]. The stable dimer was formed by incubating plasma-derived wild-type antithrombin at pH 5.7 for 16 hours. The structure revealed a domain swap of around 50 residues in which strands 4 (RCL) and 5 of β-sheet A of one subunit were inserted into β-sheet A of the next subunit. To test the relevance of the structure to AAT, cysteine mutations were introduced at adjacent positions on strands 5A and 6A so that intermolecular disulphide bonds would only form if polymerisation involved such a domain swap. Polymerisation was achieved using Gu-HCl. This experiment demonstrated a swap of strand 5A during polymerisation, as the oxidised monomer did not polymerise. However, Ekeowa *et al*. argue that 5A is not involved in the linkage as peptides with homology to strand 5A did not block and reverse polymerisation[69]. Furthermore, a distinction must be made between the contribution of introduced disulphides to overall stability and direct interference with the polymerisation mechanism[74].

The β-hairpin model for linear polymerisation as shown in Figure 1.8, was generated by breaking one of the intermolecular contacts observed in the crystal structure and placing the monomers in a head-to-tail configuration. Notably, to create a flexible intermolecular linker, helix I is proposed to unravel, which is a feature not seen in the crystal structure. This was tested using limited proteolytic cleavage by Asp-N and N-terminal sequencing, which showed that polymers were more susceptible to digestion than monomer and sites around helix I were cleaved[78]. In direct contradiction to this data, Ekeowa *et al.* found no cleavage by Lys-C and Asp-N within helix I of polymers formed by heating and from polymer isolated from hepatocytes[69]. Similarly, NMR characterisation of the Queen's AAT variant showed high stability in the helix I-s5A region[82]. As heat and liver polymers are recognised by the polymer-specific 2C1 monoclonal antibody and Gu-HCl polymers are not, this suggests a non-physiologically relevant polymer has been formed by the latter treatment. The β-hairpin model, as presented, also proposes an intermediate in which the hydrophobic β-hairpin is exposed[78]. Polymerisation-prone mutants, however, are known not to activate the UPR, which makes the presence of a grossly unfolded region unlikely[60,69]. In folding experiments, strand 5A was also shown to associate with the rest of the molecule before more C-terminal regions in contradiction to the β-hairpin polymer forming from an intermediate on the folding pathway[93]. Lastly, a study which used ion mobility MS to assess the collision cross section of the AAT dimer concluded that the theoretical loop- β sheet A dimer, but not the β-hairpin dimer, gave a collision cross section in good agreement with the experimental dimer values[69]. At the time this study was carried out, the C-terminal model had not been proposed.

E) The C-terminal model

The C-terminal model is based upon the crystal structure of a self-terminating AAT trimer solved to 3.9Å[86]. This polymer was formed from recombinant, non-glycosylated AAT that had an introduced disulphide to prevent the β-hairpin swap, by heating at 60ºC for 18 hours. It was recognised by the 2C1 monoclonal antibody, demonstrating that is shares an epitope with the pathological polymer. The structure of the circlet revealed a domain swap between subunits of the C-terminal 34 residues, encompassing strands 1C, 4B and 5B. The RCL of each monomer was self-inserted into β-sheet A, consistent with polymer hyperstability. The linkage also allows considerable flexibility in the polymer chain, in accordance with the flexibility that is

observed for polymers visualised by NS-EM[32,41,81,94]. In the crystal structure, there were no exposed unfolded regions; however, the hydrophobic C-terminus must be exposed during formation of the linear polymer model, which would be expected to activate the UPR. The domain swap is supported by experiments investigating the folding pathway of AAT: s1C/s4B/s5B associate last with the rest of the AAT fold and this is critical to keeping the RCL solvent-exposed[93]. The structure was further supported by disulphide-trapping polymerisation experiments *in vitro* and in COS-7 and *Pichia pastoris* expression systems[86]. However, the source of the material subjected to crystallography was an atypical, enriched band not normally seen in polymeric material by non-denaturing gels. This fact, combined with the unusual disulphide trapping approach, renders its relevance to a biological context unclear.

The solution structures of Z-AAT polymers formed by *Pichia Pastoris* were also investigated by SAXS[95]. The trimeric, tetrameric and pentameric polymers all produced a scattering pattern and distance distribution curve only consistent with a ring-like structure. Both comparison with the theoretical scattering curve of the crystal structure trimer and *ab initio* modelling suggested a structure similar to that observed in the crystal structure, although not as symmetrical. This, together with a modelling study excluding the loop-β sheet A mechanism on the basis of rigidity preventing the formation of small cyclic polymers[91], supports the C-terminal model of polymerisation. However, it is worth reiterating that the polymers isolated from livers and imaged by EM are mostly linear and those that form self-terminating rings are composed of greater than three subunits[32,81].

1.8.2 Drawing the molecular models together

The loop-β sheet A, β-hairpin and the C-terminal models of polymerisation are all proposed to represent the AAT polymer formed in the liver of patients with AATD[32,78,86]. However, they have all been based on polymerisation experiments or crystal structures using artificial means of polymerisation. The relevance of *in vitro* polymerisation data to the *in vivo* polymer is unclear. It is clearly possible to form AAT polymers with differing inter-molecular linkages; therefore, it is of great clinical interest to determine the linkage in *ex vivo* material and which method of artificial polymerisation yields this linkage in order to aid laboratory studies.

The models agree that β-sheet A is 6-stranded in the polymer and this is due to insertion of the RCL. However, a major difference between them is whether the RCL is self-inserted and the single subunit resembles the cleaved conformation; this is the case for the C-terminal model but not the other two models. The linkages proposed either have one, two or three β-strands that are exchanged between neighbouring subunits[32,78,86]. The exchange of three β-strands from the B and C β-sheets in the C-terminal model mean that there is only a single chain between subunits whereas for the loop-β sheet A and β-hairpin models there are two chain linkers. This has significant effect on the flexibility of these models, whereas the C-terminal model is the most flexible. This corresponds well to the flexible nature of polymers observed by NS-EM[41,81]. Furthermore, the species crystallised was generated by heating and recognised by 2C1, suggesting structural similarity to *ex vivo* material[76,86]. This contrasts to the β-hairpin species that was made by acid incubation and was not recognised by 2C1[78]. Lastly, *ex vivo* polymers have mostly been observed to be long and linear. The C-terminal and β-hairpin models were both based on atypical self-terminating structures that required modifications to form a linear model. There is not a structure of the loop-β sheet A model to make the equivalent comparisons.

The controversy regarding the polymerisation mechanism for *ex vivo* polymers and the structural similarity between *ex vivo* and heat-induced polymers has been addressed in this thesis by using monoclonal antibody technology (Chapter 4), negative-stain EM (Chapter 5) and cryo-EM (Chapter 6).

1.9 Treatment of AATD

1.9.1 Current treatment of AATD for patients

As AATD manifests in the lungs and the liver, close monitoring of, and therapeutic strategies for, both organs are required. However, there are a lack of AATD-specific therapies. The current treatment of patients is largely preventive and symptomatic, with organ transplantation suitable for end-stage disease[96]. Recommended behavioural or lifestyle changes include: smoking cessation, avoidance of environmental pollution, increased exercise, diet regulation and reduction of alcohol consumption[97]. Up-to-date vaccinations against influenza, pneumococcus and hepatitis are also recommended[97]. These actions help to preserve remaining lung and liver function. Medical therapies such as bronchodilators, corticosteroids and supplemental oxygen aim to relieve emphysema symptoms.

The only treatment that directly targets the deficiency state is augmentation therapy, which involves the delivery of purified intravenous human plasma AAT to correct the plasma deficiency[96]. The predominant therapeutic product historically has been observed to be conformationally impure[26,98]. Augmentation therapy is standard in the US and some European countries. However, evidence for lung function improvement is controversial and currently this is not a treatment option in the UK[43]. Inefficacy has been put down to a lack of suitably powered studies, as well as lung function being an inaccurate measure of disease progression[43]. This has been addressed by the recent randomised, double-blind, placebo-controlled RAPID trial. This showed that the rate of lung density loss, as assessed by CT scans, was significantly lower in the group receiving replacement AAT than in the placebo control[99]. There is also no specific treatment for the liver disease. Therefore, treatment for end-stage disease is transplantation. AATD accounts for 3.2% of adult lung transplants in the UK and 3.5% and 1.1% of paediatric and adult liver transplants, respectively[96]. A transplanted liver retains its endogenous level of AAT expression and therefore may help to prevent the progression of pulmonary disease[100,101].

1.9.2 Therapeutic approaches in development

With a greater understanding of the molecular basis underlying AATD, many therapeutic targets have been identified. These range from pathways involved in the intracellular processing of Z-AAT, sites on the Z-AAT molecule that can block intracellular polymerisation and genetic silencing or engineering of the Z mutation[96].

A) Modifiers of intracellular AAT processing

As the Z mutation causes a misfolding of AAT, chemical chaperones have been investigated as therapeutics to stabilise mutant intermediates and reverse retention within the ER. The fatty acid, 4-phenylbutyric acid (PBA), showed modest increases in the secretion of active AAT in cell and transgenic mouse models of AATD, but no significant improvement in a small study of patients with this condition[102,103].

Correcting the inappropriate trafficking of Z-AAT has been targeted using the histone deacetylase (HDAC) inhibitor, suberoylanilide hydroxyicacid (SAHA). SAHA increased functional Z-AAT secretion to 50% of the level of M-AAT secretion from epithelial cell lines[104]. The acetylation status of cellular components via HDACs and histone acetyltransferases have a role in proteostasis regulatory pathways. It is

thought that the effect of SAHA was partly due to inhibition of HDAC7 and modulation of the interaction with calnexin. Calnexin is a chaperone that interacts with AAT in the ER by recognising monoglycosylated N-linked glycans[35,105].

The stimulation of degradation pathways is an alternative approach to the clearance of intracellular AAT polymers. Autophagy is one pathway activated by AAT accumulation and involved in the degradation of insoluble polymers[106]. Very high doses of the autophagy-stimulating drug carbamazepine were effective in decreasing AAT hepatocyte polymer load in a mouse model of AATD. It is now in a phase 2 clinical trial for AATD patients[96,107]. Weekly-dosing of another autophagy-inducing agent, rapamycin also reduced intracellular accumulation of Z-AAT polymers and markers of hepatocellular injury in the AATD mouse model[108]. Finally, similar improvements in the mouse liver were observed by overexpression of the autophagy regulator, transcription factor EB, by viral vector delivery to hepatocytes[109].

B) Blocking intracellular polymerisation

Early work found that peptides homologous to the RCL could bind into β-sheet A of AAT and block polymerisation *in vitro*[32]. These peptides were optimised for potential therapeutic delivery (by shortening to four residues) and for Z-AAT specificity[110,111] . The most potent peptide (Ac-TTAI-NH_2) has also been shown to prevent and reverse Z-AAT polymerisation in cellular models[112], which could help treat Z-AAT related liver disease. All *in vitro* studies describe a loss of inhibitory activity of peptide treated Z-AAT and therefore AAT replacement would need to be administered in parallel, to treat AAT-related emphysema[110-113]. There have been no human trials reported as there is a remaining challenge of targeting the peptides to liver cells[96].

A crystal structure of native AAT helped identify alternative sites on the molecule that could be targeted by drugs. Importantly, these were distinct from β-sheet A with the aim of preserving AAT inhibitory activity[8]. One hydrophobic cavity, lined by strand 2 of β-sheet A and helices D and E, becomes occupied on transition from the native state to RCL-inserted state[8]. Its potential as a drug target was tested by a pocket-filling mutation, which reduced polymer formation and increased the secretion of Z-AAT from a Xenopus oocyte model[114]. Virtual ligand screening identified small molecules termed the 'Covent Garden (CG)' series that targeted the cavity and blocked polymerisation *in vitro* and in a cellular model[115]. These studies show the

potential to block AAT polymerisation through small molecules but CG-bound AAT was still inactive as a protease inhibitor[115].

Lastly, a monoclonal antibody named 4B12 has been developed that can reduce polymerisation and preserve AAT inhibitory activity both *in vitro* and when expressed as an intrabody in a cellular model[116]. While not seen as a viable therapeutic in its own right, characterisation of the epitope at helices A and I provides another site that could be targeted by therapeutic small molecules[117].

C) Genetic modifications

The effects of polymerisation and accumulation of Z-AAT in the liver has directed approaches to silence expression of Z-AAT by small interfering RNAs. An antisense oligonucleotide complementary to the AAT gene has been found to inhibit AAT expression in mice and monkeys[118]. As a result, there was significantly reduced AAT accumulation in the liver and reduced liver disease, with even a reversal in liver fibrosis after long term treatment. This approach has the potential for a liver disease treatment but also reduces circulating levels of AAT and so would need to be partnered with another treatment to not exacerbate lung disease.

Correction of the Z genetic defect has been achieved in differentiated human induced pluripotent stem cells (hIPSCs)[119]. Patient-specific hIPSC lines are another useful model of AATD: they have been generated from dermal fibroblasts of patients with the disease and differentiated into hepatocyte-like cells. The genetic correction was achieved using a combination of Zinc finger nucleases and a PiggyBac DNA transposon. It effectively restored wildtype AAT behaviour in the cells, including after transplantation into a mouse model of liver injury. It remains a promising approach for the future with active investigation of gene-editing tools such as CRISPR/Cas9 as therapeutic agents[96,120]. In a recent study, protein aggregation, liver fibrosis and inflammation were reduced in the liver of a gene-edited ZZ mouse model[120].

1.10 Specific Aims and Approach

The objective of this thesis was to further the structural understanding of both artificially-made and patient-derived AAT polymers. It is hoped that this will inform future therapeutic approaches that target the polymerisation process directly. The work was primarily achieved by combining the use of conformational-specific monoclonal antibodies and electron microscopy. In particular, the overarching aim

was to determine the structure of the pathological *ex vivo* AAT polymer, to a high-medium resolution (ideally around 7-10Å) using cryo-electron microscopy (cryo-EM) to reveal the polymer inter-subunit linkage. The material was extracted and purified from explant liver tissue obtained from the Birmingham BioBank, and obtainable only in very small quantities. Cryo-EM was the method of choice as it can accommodate the natural origin, high flexibility, chain length heterogeneity and glycosylation of the polymers. These characteristics make the polymers unsuitable for other structural techniques such as crystallography and NMR. A structural understanding of real pathogenic polymers will be crucial for rational drug-design, which will hopefully contribute towards the development of therapeutics that can prevent and reverse polymerisation and the associated liver and lung pathology.

To work towards the thesis objective, three branches of work were pursued that are covered in the three results chapters:

1) Development and characterisation of monoclonal antibodies against AAT that show novel properties. The monoclonal antibody 1D9 was identified from a hybridoma screen to have a strong affinity for monomeric AAT, complementing the polymer-specific antibody 2C1[76]. The properties of 1D9 have been characterised in-depth using a variety of assays and techniques, including obtaining a crystal structure of the 1D9 Fab region alone and in complex with AAT.
2) Negative-stain EM of *ex vivo* and heat-induced polymers. Heat-induced polymers share an epitope with *ex vivo* polymers and are thought to be a suitable model for the pathological polymer due to their common recognition by 2C1[76]. Heat-induced and *ex vivo* polymers have been studied bound with 4B12 Fab by NS-EM single particle analysis. Processing has been carried out with two image processing packages: IMAGIC[121] and Relion[122]. The resultant low resolution structures of dimers have been assessed and compared with molecular models of the proposed polymerisation mechanisms.
3) Cryo-EM of *ex vivo* polymers. *Ex vivo* polymers have been extracted and purified from liver samples of ZZ patients with AATD for characterisation. They have also been labelled with 4B12 antibody Fab domains and imaged by cryo-EM for the first time. Image processing using Relion[122] has yielded low-resolution structures of the dimer and the single subunit, providing a strong basis for future work.

The primary techniques used to obtain these results - monoclonal antibody technology, crystallography, NS-EM and cryo-EM - will be introduced within the subsequent chapters.

2.0 Introduction to Electron Microscopy (EM)

2.1 Introduction

The combination of cellular and structural biology over the last ~200 years has shed much insight into the mechanisms, functions and dynamics of molecules involved within a wide range of biological processes. Optical microscopy was the first technique used primarily for visualisation of cells, and later to provide information on inner cellular organisation. The invention of fluorophores made it possible to track the spatial distribution and movement of molecular complexes within cells. The techniques of X-ray crystallography and Nuclear Magnetic Resonance (NMR) have revealed detailed structures of many small proteins and their complexes with ligands. The information obtained by electron microscopy (EM) spans both levels of detail: scanning EM and transmission EM (TEM) of cellular sections or tomography can provide cellular details, while single particle analysis EM can achieve medium-high resolution and reveal different conformations of proteins and complexes.

2.2 Optical microscopy in biology

The human eye can resolve details up to ~0.1mm in size. Optical microscopy was invented to extend this limit and was based on the use of lenses to magnify images of objects and therefore improve the visibility of details. The first optical microscopes, pioneered by Robert Hooke and Antoni van Leuwenhoek in the 17th century, were able to provide magnified images of objects up to 250 times and could resolve 1-10μm details[123]. With the development of compound microscopes that use complex lenses, this is now at 1500 times magnification, resolving details of 200nm in size. Better visualisation of cellular components was also aided by selective chemical or fluorescence staining, overcoming the inherently low contrast within biological specimens.

The assessment of a resolution limit in the light microscope was given by Ernst Abbe, who has shown that the resolution (d) can be defined as the smallest observable distance between two separated features. It depends on the wavelength of the light used in the microscope (the wavelength of visible light is in range of 200-750nm) and the aperture of a microscope. The physical limit is defined by the Rayleigh criterion:

$$d = \frac{0.61\lambda}{NA}$$

where NA is the numerical aperture defined as NA = nsinθ, where n is the refractive index of the medium and θ is half the acceptance angle of the lens/aperture (Figure 2.1A). The criterion originates from the scattering of the illumination wave after it interacts with an object. An Airy disc with a central bright region and a series of surrounding concentric rings, rather than a perfect circle, is formed in the image plane when light passes through a circular aperture[124]. This Airy disc represents a point spread function (PSF) of the optical system. The resulting image of the object is a convolution of the PSF with the ideal image. Two points are described as resolved when the first minimum of the Airy disk from one point falls on the central maximum of the other (Figure 2.1)[124]. Super-resolution fluorescence microscopy has surpassed the limit described by Abbe, by reducing the PSF dimensions and reaching resolutions of ~50nm[125].

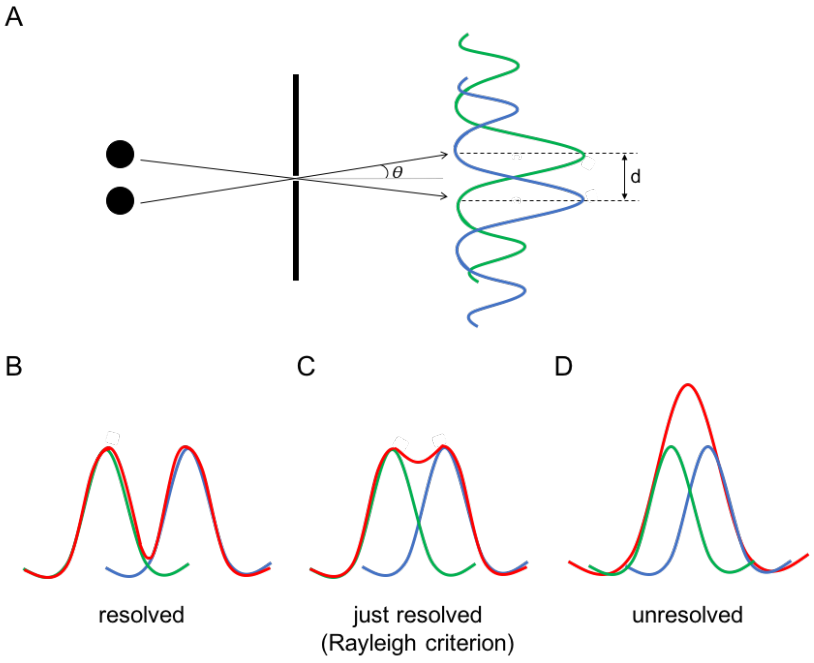


Figure 2.1. The Rayleigh criterion. *A) The Rayleigh criterion for resolving two points. B) Two points (blue and green curves) appear as distinct intensity peaks in the image when they are far apart. C) Two points are just resolved when they are distance d apart, where the minima of one PSF overlaps with the maxima of another. D) Two points are unresolved when they are close together such that their PSFs overlap to form one maximum (red line).*

2.3 High resolution molecular techniques

X-ray crystallography and Nuclear Magnetic Resonance (NMR) have structurally characterised thousands of proteins at an atomic resolution (1-4Å). X-rays have wavelengths of 1Å, close to the size of atoms. A combination of diffraction patterns collected at different angles, with phase information, corresponds to the 3D Fourier transforms of the molecular complexes from which the crystal is formed. The reversed Fourier transform generates electron density maps of the crystal repeating unit and these densities are used for model building (more details in Chapter 4.1.5). In contrast, NMR uses radiofrequency waves with wavelengths of centimetres-metres to irradiate a sample and excite certain nuclei. The radiation emitted by the nuclei on relaxation is recorded and is specific to the structural environment of the nuclei. Information about the nuclei in a protein sample provides multiple structural constraints which can be combined with modelling and used to obtain atomic resolution structures. Despite their achievements, both techniques have serious limitations. X-ray crystallography depends on the growth of well-diffracting protein crystals and in NMR the spectra of proteins greater than 50kDa are difficult to interpret due to the overlap of nuclei peaks. These problems are generally exacerbated for large proteins, which means that protein complexes are more inaccessible to structural studies by these methods.

TEM is uniquely suited to the study of structures greater than 100kDa and is less demanding in sample amount and homogeneity. This has extended the range of protein complexes amenable to structural determination and together with the technical advances, has offered an alternative route to high-resolution structure determination.

2.4 Imaging using electrons

According to the theory of dualism of particles formulated by Louis de Broglie in 1922, electrons can also be considered as waves. Accordingly, a particle with a mass (m) and that moves with a velocity (v), can be scattered by matter and diffract according to the principles of conventional light optics. A wavelength (λ) of the particle is defined by:

$$\lambda = \frac{h}{mv}$$

where $\hbar$ is the Planck constant. Since electrons are charged particles they can also be deflected and focused by an electromagnetic field. This principle has been used in construction of the electromagnetic lenses for the first transmission electron microscope by Ernst Ruska in the 1930s[126].

The short wavelength of accelerated electrons make them suitable for high-resolution structural imaging. The wavelength of electrons depends on their kinetic energy (eV = ½ $mv_{2,}$ where e is charge and V is the acceleration voltage) and can be defined by:

$$\lambda = \frac{1.22}{V^{1/2}}$$

For an accelerating voltage of 100kV and 300kV, the wavelength of the electrons is equal to 3.9pm and 2.2pm respectively. Then according to the Rayleigh criterion, the achievable resolution attainable in EM using a numerical aperture of 0.009 are 0.26nm and 0.15nm for these two wavelengths. However, for biological specimens the realistic attainable resolutions are much worse due to sample preparation difficulties, low image contrast and signal to noise limitations. Still, in recent years due to technological advances in hardware and software, the resolution of EM reconstructions has improved up to 1.8Å[127].

The highest resolution electron density maps obtained in TEM are using the single-particle analysis approach. Here, thousands of randomly oriented particles are imaged; these images represent 2D projections of individual molecules. Particle images that are similar to each other are grouped in classes and averaged to improve the signal to noise ratio (SNR). The orientations of each 2D class are then estimated and these images are used to reconstruct an initial 3D volume. The other main approach which yields lower resolution structures, but can be informative in terms of cellular structure and in-situ spatial organisation of particles, is electron tomography. This technique records multiple 2D projections of the same field of view at different angles to allow reconstruction of a 3D volume.

2.5 The electron microscope

An electron microscope consists of an electron source, electromagnetic lenses to make the electron beam parallel when it illuminates a specimen and then to focus the scattered electrons on a detection system where the magnified image is formed (Figure 2.2).

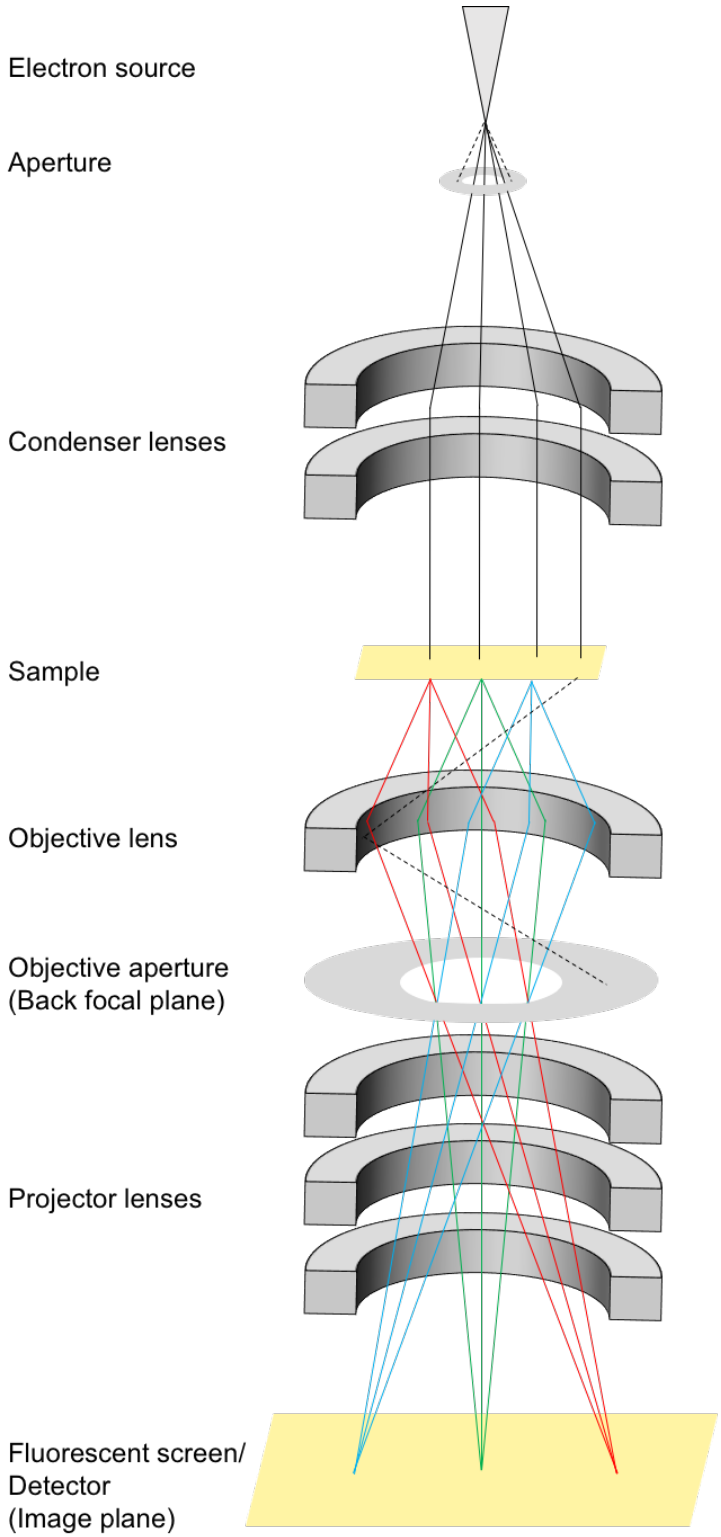


Figure 2.2. Scheme of an electron microscope. *The condenser lenses parallelise the electron beam from the source to illuminate the sample. Electrons are elastically scattered (shown in red, blue, green) by the sample. High-angle scattered rays are blocked by the objective aperture. The projector lenses focus the magnified image onto a detector system.*

2.5.1 Electron sources

The electron source is located at the top of the microscope and produces the electrons used for imaging. The electron beam needs to be temporally coherent (all electrons are emitted with the same energy) and spatially coherent (emitted from the smallest area)[128]. This reduces the effect of lens aberrations and improves the achievable resolution of the microscope. The beam should also be intense and bright.

Traditional electron sources are thermionic where electrons are emitted from a metal such as a tungsten filament, which is heated to temperatures around 2000-3000℃. The emitted electrons are then accelerated by an electric field towards an anode. LaB_6 crystal sources have finer tips that require a lower temperature for emitting electrons and provide a more coherent beam. A much more parallel, brighter and monochromatic beam is produced by a field emission gun (FEG). Electrons are extracted from a single tungsten crystal, that is sharpened to a thin needle, by an electrostatic field and further accelerated by a voltage of 100-300 kV.

2.5.2 The lens system

Electron microscopes have a complex system of 5-6 lenses. The electromagnetic lenses used in EM are solenoid coils that can deflect an electron and whose power is controlled by varying the current flow within the coils. The first lenses in the optical system of the electron microscope are the condenser lenses, which create a parallel beam of electrons that irradiate a sample. The specimen is located just above the objective lens. The objective lens provides the primary magnification (20-50X) and hence its aberrations have a large impact on acquired images[129]. In the back focal plane of the objective lens the Fourier transform of the sample image is formed (described later). The objective aperture is located in this plane. The size of the aperture defines the angle at which electrons scattered by the sample will be cut off from formation of the image by projector lenses and is used to increase amplitude contrast in the image. A set of 2-3 projector lenses, located below the objective lens, provide the final magnification of the TEM image onto the phosphor screen or a recording device of the microscope.

The quality of EM images depends on aberrations of the lenses. The most common aberrations are spherical, chromatic and astigmatism (Figure 2.3)[129]. In a perfect

lens, any parallel beam will be focused in one point in the back focal plane of the lens. Spherical aberrations of lens are related to the different focal distances for the rays that are passing at the periphery or through the central part of the lens. The rays that passed through the periphery are focused more strongly than rays closer to the optical axis. This means the focal length varies with distance from the optical axis. For chromatic aberrations, the focal length varies with the wavelength of the incident electron wave so that shorter wavelength electrons are focused closer to the lens. Astigmatism results from an asymmetrical magnetic field in the lens whose focusing power differs in perpendicular directions. It presents as elongation of features in one direction or an elliptical shape of a spherical point.

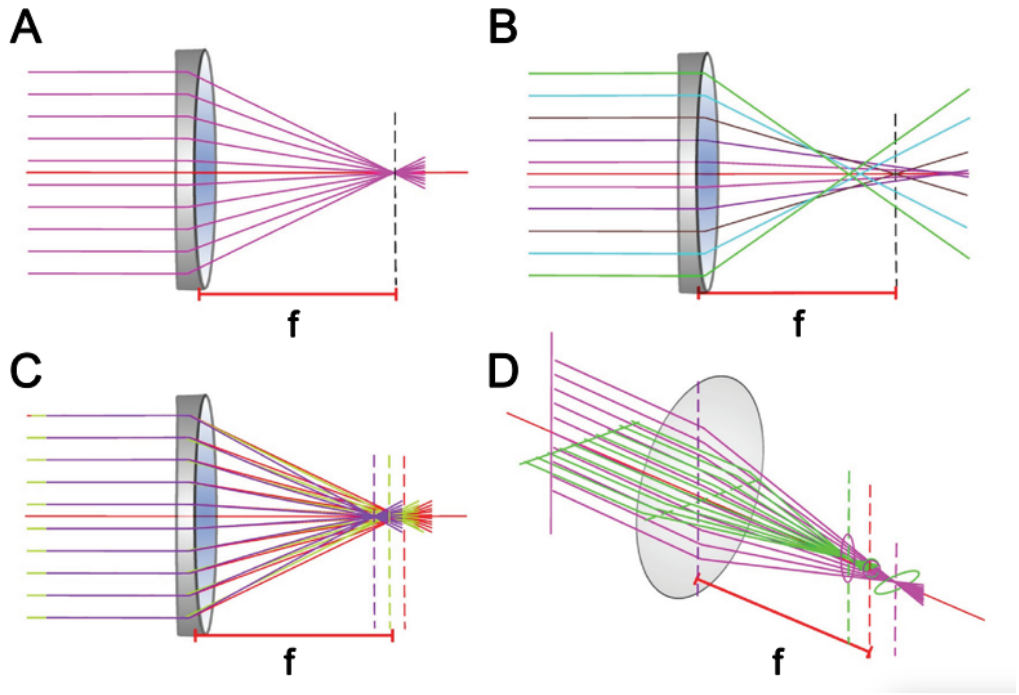


Figure 2.3. Schematics depicting the aberrations of lenses. *A) a perfect lens, B) spherical aberration, C) chromatic aberration, D) astigmatism. f is the focal length of the lens. Figure reproduced with permission from* [129].

2.5.3 Detectors

The first TEM images were recorded on photographic film. This was cheap, with a large field of view and fine spatial resolution (pixel size of 6 μm). However, all micrographs had to be manually scanned for subsequent digital image processing

meaning the response time was slow and this limited the ability to collect high-throughput TEM data.

Film was replaced by charge-coupled device (CCD) cameras. These detectors rely on sensors that have two steps: the conversion of electrons into photons by a phosphorous scintillator and then the conversion of photons into charge via fibre coupling. A camera has an array of sensors that comprise the CCD chip. However, for image readout, the charges are serially transferred to a readout register and digitised. These cameras have several disadvantages. First, they are less sensitive at voltages above 120 kV. This is because high energy electrons have a reduced scattering cross-section and so thicker scintillators are needed to maintain the yield of photons. An additional problem is that secondary photons produced by electrons scattered within the phosphor layer, lead to registration by several adjacent sensors and increasing of the PSF size and noise in the image, reducing the resolution[128]. CCD sensors (pixels) are large, about 24 μm, meaning that a higher magnification is needed to achieve the same pixel size and thus reducing the field of view.

High resolution TEM has become more accessible largely due to the advent of direct electron detectors (DEDs). DEDs consist of small (5-6 μm) sensors that can directly register electrons[130,131]. Due to the absence of an energy conversion step from electron to photon (where significant energy loss takes place) and shorter detection distance, DEDs offer a fast readout, a smaller PSF and improved SNR compared to CCDs. This also means they can rapidly record multiple subframes per image in the "movie mode" to allow for correction of sample drift, as well as finer control over the total electron dose that the sample has been exposed to in the images used for reconstruction[132]. They are more expensive and produce datasets that require greater disk space for storage and processing, however their advantages for high-resolution structural determination are clear.

2.6 Usage of Fourier transform in EM

Within the microscope, a Fourier transform (FT) is formed on the back focal plane of the objective lens. A FT is the decomposition of a complex function into a series of sine waves[128]. Representation of a function as waves is often described as being the "reciprocal" space. Reciprocal (or Fourier) space and real space (the space of the original function) can be readily interchanged by a Fourier transformation. The reversed FT takes place in the microscope as the scattered wave moves through the

projector lenses towards the image plane, where the component waves re-interfere forming the image that is registered by detectors.

Fourier decompositions or summations may not always be exact as it depends on the highest frequency of wave used to restore the original function[128]. The Nyquist frequency is the highest frequency component present in the calculated FT of the function. This is the sine wave that oscillates with a frequency half that of the sampling rate (i.e. the pixel size divided by 2)[128]. The high frequency components are represented near the edge of the FT, whereas the low frequency components (and low-resolution information) are located near the origin.

Procedures in image processing can be carried out in real or Fourier space. There are fast computer algorithms to calculate the FT and convert between them. Furthermore, calculations are often faster in Fourier space, as it converts differential operators in real space into multiplication operators.

2.7 EM sample preparation for single particle analysis

2.7.1 Background

The electron path is kept under high vacuum in the microscope. This is because electrons interact strongly with matter and otherwise their path would be impeded by gas particles. The vacuum requires that samples to be used for EM imaging must be in the solid state as any moisture from biological specimens would be evaporated causing structural collapse. Consequently, samples must be dehydrated or stabilised in a hydrated state before imaging. The chosen method for this depends on the sample and desired resolution of reconstruction.

2.7.2 Negative staining

Negative staining (NS) is the quickest and simplest method for preparing grids for TEM studies. It can be used to obtain low-resolution structures or for assessment of sample purification and homogeneity before cryo-EM.

NS fixation involves the application of a heavy metal salt solution to the EM grid after application of the sample. Excess liquid is blotted away and the grid is left to air dry. The heavy metal atoms surround the sample molecules on the grid, creating a thin layer of dark granular background. Common stains are based on tungsten and

uranium, for example a 1-2% w/v aqueous solution of uranium acetate (UA). Although UA offers the best contrast, its acidic pH (~4) can lead to structural changes or complex disassembly and its microcrystalline structure is sensitive to the electron beam. This makes phosphor-tungsten acid a popular choice with its more physiological pH.

NS-EM produces a 3D envelope of the molecule that reflects the shape and solvent-exposed surface, but the obtainable resolution is limited to ~15-20Å due to the grain size of the metal salts[133]. A further consideration is that NS can also result in distortion of molecular shape (e.g. flattening) and dehydration can cause the structure to collapse[124].

2.7.3 Cryo-fixation

Cryo-fixation allows higher resolution of structures in a more native-like state to be achieved due to the absence of stain. The sample suspended on an EM grid is rapidly cooled to liquid nitrogen temperatures (77K) by plunging into liquid ethane. The rapid drop in temperature preserves the water as a vitreous, non-crystalline solid state preventing the formation of hexagonal or cubic ice, which would damage biological material. Ethane must be used as it has a high heat capacity and a high heat conductivity providing extremely fast cooling of grids[129]. The liquid ethane is maintained at its freezing temperature by liquid nitrogen. Ethane is more efficient at retrieving heat from the sample than liquid nitrogen and does not boil as easily, therefore preventing an insulating gas layer from forming around the grid.

The specimen is retained in a thin layer of vitreous ice as a snapshot of a solution-like state. The contrast of cryo-EM samples is much lower than in NS-EM as the atoms composing biological macromolecules are very like those of water and buffer components. A hybrid approach is cryo-negative EM which can prevent dehydration and specimen collapse and increase contrast, however the attainable resolution is limited by the grain size of negative stain salt[134,135]. The typical stain used here is ammonium molybdate. Plunge-frozen samples must be maintained at cryogenic temperatures from sample preparation to grid insertion into the microscope and imaging.

2.8 Image formation

2.8.1 Interaction of electrons with the specimen

There are several different outcomes from the primary electrons hitting the sample[129]. Some electrons pass through the sample without any interaction (unscattered). Others are deflected by the positive electrostatic field of the nucleus or the negative orbital electrons of specimen atoms, and some are scattered at high angles or backscattered after collision with atomic nuclei or electrons.

The interactions of electrons with the sample can be classified as elastic or inelastic scattering. An elastic scattering event does not change the kinetic energy of the electron. Instead, it is deflected by an angle proportional to the nuclear charge and inversely proportional to its energy and distance from the nucleus. An electron involved in an inelastic scattering event transfers or gains some energy from the specimen. This has two major consequences: radiation damage to the specimen and the electron of altered energy being focused at a different point, leading to the blurring of the final image. The objective lens aperture and an energy filter can remove energy-loss electrons from the final image. Damage to the specimen structure can be through the ionisation of atoms, induction of X-ray emission, chemical bond rearrangement, generation of free radicals or causing secondary electron scattering[136]. Radiation damage often limits high-resolution structural determination as features of biological specimen can be affected at electron exposures as little as 10e/Å^2. To alleviate the damage in recorded images, low dose modes (as well the recording of movies, discussed later) were introduced into imaging software. These use a deflection of the electron beam to a neighbouring region for area selection and focusing before image acquisition.

2.8.2 Origins of image contrast

Contrast in TEM micrographs is defined as the difference between the brightest and darkest pixel intensities divided by the average density in the image[129]. There are two types of contrast in biological TEM: amplitude and phase contrast (Figure 2.4).

Amplitude contrast arises in the image from absorption of some electrons from the incident beam. Atoms with high atomic numbers (e.g. heavy metal stains) absorb or scatter electrons at high angles. These will not contribute to the final image; the latter

as it is blocked by the objective lens aperture. Such regions in the specimen will appear dark in the image. Biological specimens however produce little amplitude contrast as they consist of light atoms that do not absorb electrons. In general, the total number of electrons before and after the specimen remains nearly the same. However, the angle (and consequently the path length) of the scattered electrons relative to the unscattered electrons changes in a location-dependent manner. Considering electrons as waves, the angle change equates to a change in the phase of the scattered electrons and creates phase contrast[128]. Phase variations cannot be directly recorded in images and so are converted into detectable amplitude variations. This is done by introducing an additional 90º phase shift between the scattered and unscattered electrons using the objective lens aberrations (Figure 2.3) and underfocussing (Figure 2.4)[128]. Underfocusing or defocusing is the reduction in strength of the objective lens such that the electrons come to a point below the image plane.

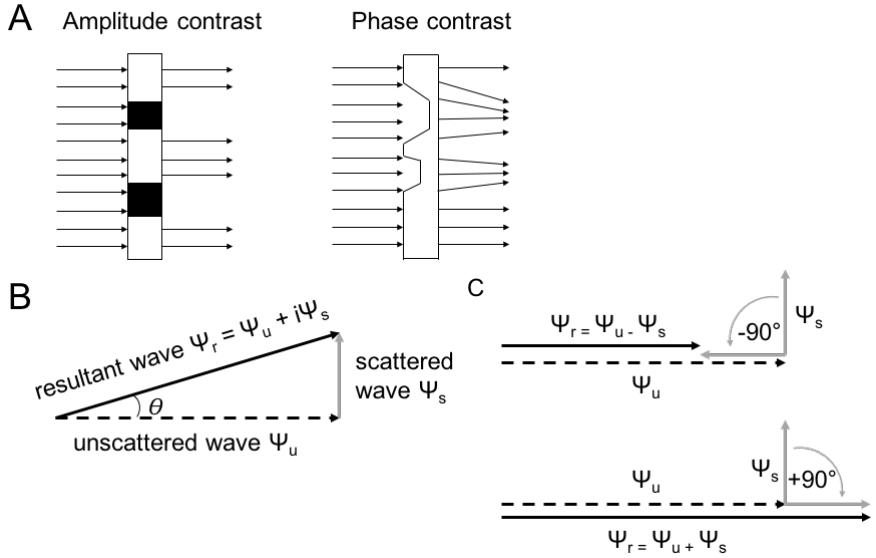


Figure 2.4. Image contrast. *A) Amplitude contrast results from absorption of some illuminating rays whereas phase contrast results from a change in direction of scattered rays. B) Waves can be represented as vectors where the vector length corresponds to amplitude and the angle θ corresponds to phase. The resultant wave has a different phase to the unscattered wave but they have similar amplitudes resulting in little amplitude contrast. C) A ±90° phase shift of the scattered wave leads to a larger difference in the amplitudes between the unscattered and resultant wave. This effectively converts the phase contrast into detectable amplitude contrast. Figure adapted from [129].*

2.8.3 The Contrast Transfer Function

Defocussing the objective lens, combined with its inherent spherical aberration, produces higher contrast, which is necessary for visualisation of biological samples. However, defocussing and aberrations also generate distortions in the image, as each point of the specimen is represented as a small blurred spot. Therefore, the recorded image is not a direct representation of the real structure due to a convolution between this blurring function (the PSF, mentioned earlier), and the real image. The convolution can be described in Fourier space, where the FT of the PSF is the contrast transfer function (CTF):

$$FT\{\Psi_{obs}(\vec{r})\} = FT\{\Psi_{ideal}(\vec{r})\} . FT\{PSF(\vec{r})\} = FT\{\Psi_{ideal}(\vec{r})\} . CTF(\vec{R}) . E(\vec{R})$$

where Ψ_{obs} is the recorded image wave function, $\vec{r}$ is a vector in the image plane, Ψ_{ideal} is an ideal projection of the sample, $\vec{R}$ is the spatial frequency, CTF is the phase contrast transfer function and E is an envelope damping function. The envelope function describes the falloff of image signal at high spatial frequencies and is caused primarily by radiation damage.

The CTF is an oscillating function that describes how the contributions of amplitude (cos term) and phase contrast (sin term) to the final image depends on spatial frequency[128]. It is defined by:

$$CTF(\vec{R}) = -(\sqrt{1-Q^2}\ \sin(\gamma(\vec{R})) + Q\ \cos(\gamma(\vec{R})))$$

where Q is the amplitude contrast ratio (~4% for 300kV cryo images[129]) and $\gamma(\vec{R})$ is the phase shift caused by defocussing and aberrations at spatial frequency $\vec{R}$ where

$$\gamma(\vec{R}) = \pi\lambda\vec{R}^2(\Delta z - \frac{C_s\lambda^2\vec{R}^2}{2})$$

where $\boldsymbol{\lambda}$ is the electron wavelength, Δz is the defocus and C_s is the spherical aberration of the objective lens. Figure 2.5 shows example CTF plots overlaid on the FT of images. The CTF reverses the phases of alternate dark and light rights (thon rings) in the FT meaning different features of the object in real space are enhanced or suppressed. Therefore, micrographs need to be corrected for the CTF before image processing (discussed later). Furthermore, where the oscillating sin/cos function crosses zero, the recorded image does not have any information at these

spatial frequencies. This information needs to be collected by recording other images at different defoci (which will have signal at the needed frequencies and have the 'zeros' at different frequencies).

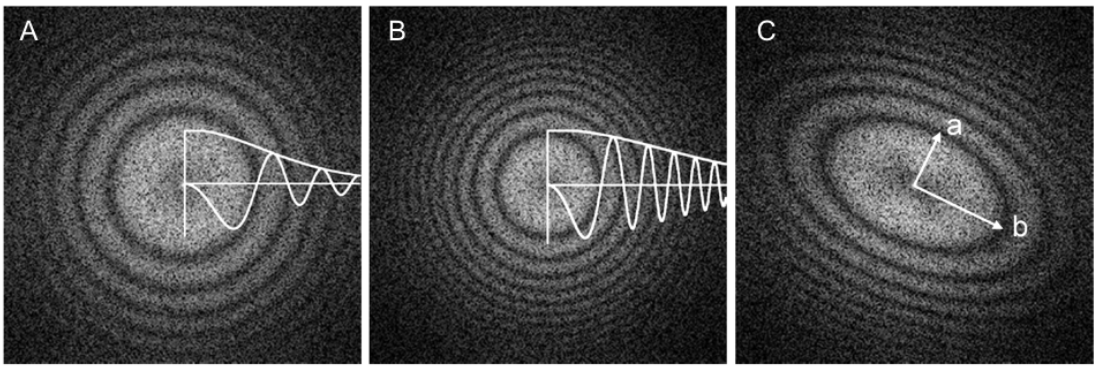


Figure 2.5. Example Fourier transforms of images. *Alternating light and dark rings (Thon rings) from the oscillating CTF curves can be seen. Image A was taken closer to focus than Image B, indicated by the Thon rings oscillating slower and the first zero being further away from the origin. The CTF plots are overlaid showing variations in the contrast contributed by different spatial frequencies. Dark regions between the Thon rings correspond to CTF nodes, where no information is collected at this spatial frequency. The amplitudes are attenuated by the envelope function. Image C is astigmatic and is indicated by the elongation of Thon rings in one direction: there is a different defocus value for the directions a and b. Figure reproduced from* [129].

2.9 Image processing – Single particle analysis

2.9.1 Pre-processing of micrographs

Micrographs recorded using DEDs have multiple subframes. The first step carried out during image processing is the alignment of the subframes[137]. This results in a motion-corrected total sum micrograph, reducing the blurriness of a solely averaged micrograph. The alignment of subframes allows for the correction of beam-induced motion and drift of the sample during an acquisition. The most significant movement is induced by the impact of the electron beam hitting the sample[137].

The averaged micrograph then needs to be corrected for the effect of the microscope CTF. As described previously some features of the image (at specific spatial frequencies) will have inverted contrast due to oscillations of the CTF. For complete restoration of the image data, the CTF needs to be corrected for and as the accelerating voltage and spherical aberration of the microscope are constant and

known, this means the defocus of each micrograph has to be established. Determination of the CTF is done by matching the experimental amplitude profile of a micrograph to a library of theoretical CTF profiles at different defoci[129]. A match allows the computational flipping of the phases of the correct negative-contrast frequencies. For each micrograph, there is no information present at spatial frequencies where the CTF crosses zero. However, during collection of a dataset, micrographs are taken are different defoci, so information at these frequencies will be retrieved from other images and all spatial frequencies will be sampled. More advanced image restoration includes the correction for amplitude decay at higher spatial frequencies due to the envelope function. These can be boosted to help achieve higher resolution structures, however a Wiener filter should also be included within the amplitude correction to avoid the amplification of noise at frequencies near CTF zeros[129].

After CTF correction, the micrographs must be normalised by scaling their densities to the same standard deviation and the mean value of zero. This is because images will have variations in contrast and grey levels in different areas of the grid and this could lead to misalignment[138].

2.9.2 Particle picking

After pre-processing of the micrographs, individual particles are selected and extracted from the micrographs. This used to be a manual process which was time-consuming but has been made easier through automated particle picking programs[139,140]. These use template matching to select particles from micrographs based on an input model or generated references. Typically, users still manually pick particles from a couple of micrographs, followed by alignment and classification to obtain 2D references (importantly the references should represent different views of the particle), which can be used as the templates for automated picking[141]. Automated programs allow control of the distance between picked particles, the threshold to select a particle and the level of filtering of the reference or micrograph. It is important to not use high-resolution references for automated picking to avoid picking ‘particles’ that are random noise[142,143].

2.9.3 Image alignment

The noise level in images is high due to the imaging of specimens with low electron doses, radiation damage, ice variations and the imperfections of the microscope/detector. Averaging of images of the same molecules in the same orientation cancels out random background noise and enhances true specimen signals. Extracted particles should be aligned and sorted into similar views or orientations of the molecules. Classes of particles representing the same view are averaged to improve the SNR by a factor of √N where N is the number of averaged images and can be used for the initial reconstruction[129].

The similarity of images is measured using cross-correlation. This scoring method underlies the alignment procedure. The value provides an indication of image similarity and the peak position of the cross-correlation function determines the optimal shift between the compared images.

Particles images are first centred by translational alignment to a rotational average of their total sum. They are then coarsely aligned, focusing on low-resolution features such as particle shape. This is achieved by low-pass filtering of the images. High frequency information can gradually be reintroduced in an iterative process, as the alignment improves and the shifts are restricted. Centred particle images can also be masked by setting all densities outside the mask to zero and hence removing background noise. A soft mask must be used to avoid hard-edge artefacts, such as ripples, in the image[129].

2.9.4 Classification of images

Classification methods are used to identify and cluster particle images into groups of identical views. In electron microscopy, image processing packages either use multivariate statistical analysis (MSA) (also known as principal component analysis) or K-means clustering combined with maximum likelihood (ML) to classify images. MSA allows the dimensionality of the dataset to be reduced and identifies the major differences between images (principal components) (Figure 2.6)[144]. Each image of m x n pixels is represented as a point in an mn-dimensional hyperspace according to its pixel density values. Similar images correspond to the points located close to each other within the cloud of images. The largest variation within the dataset is identified by the largest distances within the cloud in the hyperspace coordinate system, so a

new system should be reoriented using the directions of biggest variations as a new axis. Further major components define new axes, which are orthogonal and therefore uncorrelated and independent to the previous. Identification of the major principal components (also called eigenvectors) simplifies the analysis of the large dataset, with smaller variations attributed to noise. In image processing, eigenvectors are represented as eigenimages, which indicate the density variability of pixels in the images. Images that are localised in the hyperspace are then clustered into a specified number of classes. This is achieved either by hierarchical ascending clustering, where each image starts out defining a class and similar images are merged together or divisive clustering, where all images initially belong to one cluster and are then separated into smaller groups based on image differences[129].

The other approach is K-means clustering. In this method, a specified number (K) of random seed images are chosen to define the classes and the other images are systematically compared to the seeds and assigned to the most similar. The seeds are redefined by averaging of the class members and the process is iterated until the differences between classes are maximised and differences within a class are minimised. K-means clustering has been combined with the ML method[145], where instead of an image being assigned to a single class, there is an estimate calculated of how likely it is that the image belongs to that cluster. This means that a single image belongs to multiple clusters with different weights, until the final iteration when the classification is definitive. ML methods can also be used in the generation of multiple 3D reconstructions from a single dataset to help sort particle heterogeneity for example the existence of different conformational states[146].

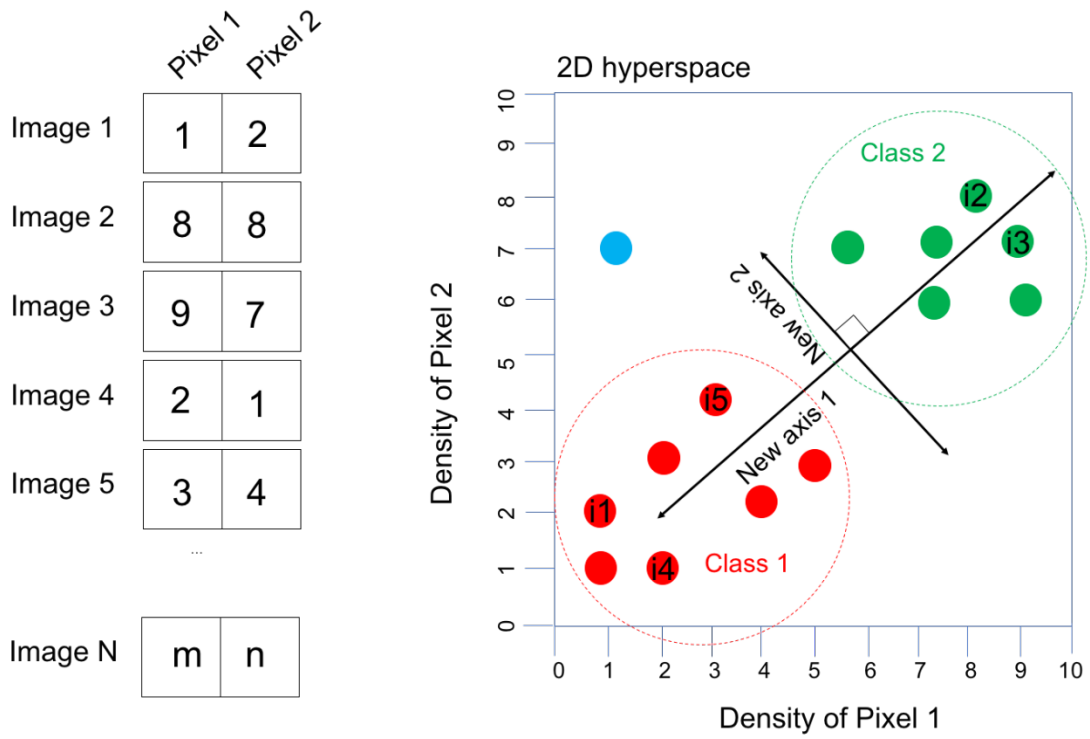


Figure 2.6. A simplified schematic illustrating the basis of multivariate statistical analysis (MSA) and classification. *A set of two-pixel images (i) are subjected to MSA and are represented as points in a two-dimensional hyperspace according to the densities of pixel 1 and 2. The two main differences between images are represented by two orthogonal, uncorrelated eigenvectors. The points are sorted into two classes (red and green) of similar images according to position in the hyperspace. An outlying point (blue) remains unclassified.*

2.9.5 Orientation determination

Using multiple different 2D projection views, a 3D electron density can be reconstructed. To do this, the views are assigned space angles using one of three methods: random conical tilt[147], projection matching[148] or angular reconstitution[149].

The random conical tilt method that provides orientations of particle images is the experimental approach[129,147]. Tilt pair images are taken during data collection, where an area of the EM grid is imaged twice at a predefined rotation about the axis perpendicular to the electron beam. The untilted particles are aligned and classified into similar views. The similar particles are then located in the tilted micrograph. As the angle of tilt is known and the in-plane rotations/shifts have been calculated from the untilted micrograph, the orientation of the particles in space is defined.

Projection matching is a computational approach for orientation determination[148]. It requires an initial model to be projected in multiple orientations, followed by the systematic comparison of the reprojections (used as a references) to the experimental particle images or class averages. The highest cross-correlation between reprojection and particle image indicates their consistency and the angles of that reprojection are assigned to the image. A new model is generated using the images with assigned angles and the process of the angular assignment iterated with a new set of reprojections until the particles/classes and model converges. Projection matching can also be carried out in reciprocal space, where 2D FTs of projections and model reprojections are compared[150].

Angular reconstitution is an alternative computational approach that identifies 'common' lines (1D projections) between 2D projections[149,151]. This method can also be carried out in reciprocal or real space. In reciprocal space, the FTs of 2D projections are planes that pass through the centre of the 3D FT of the object. This is known as the central section theorem (Figure 2.7). The line of intersection between two 2D central sections of FTs is the common line between the two FTs of 2D projections. A third 2D projection (and its FT) is needed to determine the angle between the two intersecting planes and the orientations of the projections[129]. In real space, a 1D projection is calculated by summation of pixels from a 2D projection along different directions[151]. The 1D projections from a set of 2D projections are then compared to each other by cross-correlation to find the common line between projections.

2.9.6 3D reconstruction

After angular assignment, the 2D projections are used to calculate the 3D structure of the specimen. Reconstructions can also be calculated in real or Fourier space. In real space, the back-projection method is used where 2D projections are stretched in the projection direction along the volume to be reconstructed[129]. Each pixel in the 2D projection effectively becomes a ray in the 3D volume after stretching. The density of each pixel in the 3D reconstruction is estimated by the summation of rays that pass through it. However, the 3D reconstruction is limited due to a low frequency background artefact caused by the stretching of projections. Background artefacts can be mitigated by filtering, either of the 2D projections prior to back-projection or the 3D object after back-projection (known as the filtered back-projection method)[152].

Algebraic methods are another real space approach to 3D reconstruction[153]. These use 2D projections to approximate the densities of pixels in 3D space. The newly created 3D model is reprojected and compared to the experimental projections. The comparisons are used to make corrections to the calculated 2D projections and improve the 3D reconstruction approximation. The processes of approximation and correction are iterated and can be achieved in two ways[129,153]. New projections can be sequentially incorporated into the approximation, the differences between the projection and reprojection calculated and corrections made to the previous approximation. Alternatively, differences between all projections and reprojections can be simultaneously assessed to make corrections to the 3D approximation.

The Fourier approach to 3D reconstruction is based on the central section theorem (Figure 2.7)[129]. Three-dimensional Fourier space is filled with the FTs of multiple 2D projections. The central sections do not cover all the pixels in Fourier space. Therefore, there is missing information between them and the data must be interpolated between known pixels. The inverse FT is taken of the 3D FT, yielding the electron density distribution of the 3D object[154].

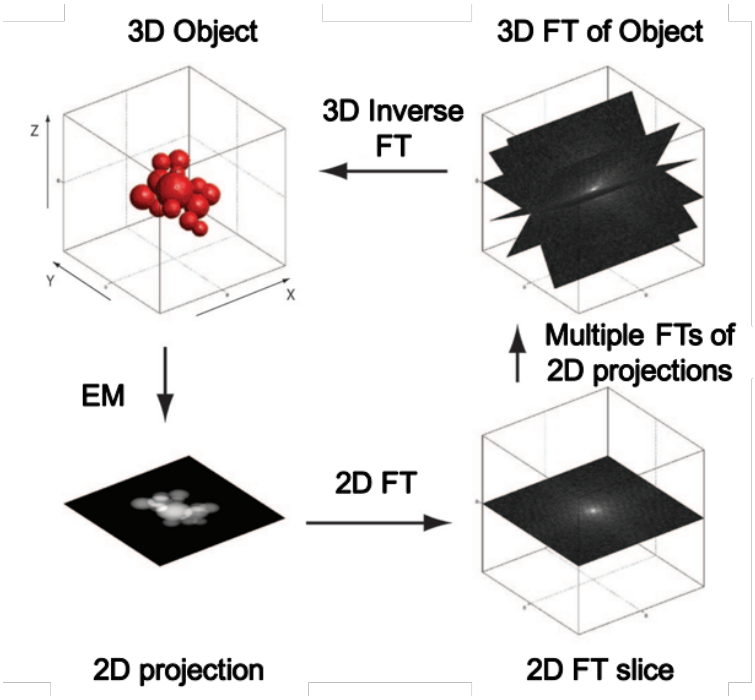


Figure 2.7. 3D reconstruction using the central section theorem. *FT denotes a Fourier transform. Figure adapted from[155].*

The angular distribution of images can affect the quality of the reconstruction. This distribution is represented by points on the Euler sphere, where the points denote the view directions of the object. Ideally, 2D projections used in a 3D reconstruction should be evenly sampled across the Euler sphere and the resulting reconstruction is isotropic in resolution. In practice, samples exhibit preferential orientations on the grid and the reconstruction can suffer from the absence of information in certain directions[129]. To rectify this experimentally, grid supports, buffer components and treatment of support films (e.g. positive glow-discharge) can be varied[156]. Computationally, the impact of projections on the final structure can also be weighted (for example higher weighting of orientations that are less populated).

Resolution of the 3D reconstruction is determined by assessing the Fourier Shell Correlation (FSC)[157]. This is useful to evaluate the quality of the structure and identify at what level the structure can be reliably interpreted. FSC assessment involves division of the dataset into two random subsets, calculation of separate 3D reconstructions and their comparison by cross-correlation in spatial frequency shells. Correlation is assessed against spatial frequency, which falls from 1 at low resolutions to 0 at high resolutions. The spatial frequency at the threshold of 0.5 correlation or 0.143 is used as the global resolution estimate[157,158]. It is also important to reproject the model at the same angles as the assigned 2D class averages/images and check for their similarity to ensure correct angular assignment. However, the most robust method is by assessing structural features using atomic models for fitting in the map. Local resolution of maps can also be assessed using software such as Resmap[159], which can reflect structural flexibility.

2.10 Interpretations

The level to which EM density maps can be interpreted largely depends on the attained resolution. Often, atomic models are docked as rigid bodies into the maps to aid interpretation. Low resolution maps (15-30Å) are useful in the definition of domains and protein interfaces in complexes. At resolutions below 10Å α-helices can be identified as rods and below 4.5Å, β-strands become apparent. This level of structural detail can be useful to aid the designing of experiments and allows flexible fitting of models. At high resolutions below 3Å, the polypeptide can be traced and side chains may be resolved for building of an atomic model.

3.0 General methods

Listed below are methods that are applicable to more than one results chapter. Methods specific to chapters are incorporated within the relevant results chapter.

3.1 List of manufacturers

Agar Scientific	Stansted, UK	Chemicals and reagents
Agilent Technologies	Santa Clara, CA, USA	Chemicals and reagents
Electron Microscopy Sciences	Hatfield, PA, USA	EM grids
Expedeon	San Diego, CA, USA	Chemicals and reagents
GE Healthcare	Chicago, IL, USA	Chemicals and reagents
Greiner Bio-One	Kremsmünster, Austria	Chemicals and reagents
Hampton research	Aliso Viejo, CA, USA	Crystallisation screens
LI-COR Biosciences	Lincoln, NE, USA	Chemicals and reagents
Life Technologies	Carlsbad, CA, USA	Chemicals and reagents
Millipore	Burlington, MA, USA	Chemicals and reagents
Molecular Dimensions	Newmarket, UK	Crystallisation screens
NewEngland Biolabs	Ipswich, MA, USA	Chemicals and reagents
Qiagen	Hilden, Germany	Chemicals and reagents
Quantifoil	Großlöbichau, Germany	EM grids
Roche	Basel, Switzerland	Chemicals and reagents
Sigma-Aldrich	St. Louis, MO, USA	Chemicals and reagents
Takara Bio Inc.	Kusatsu, Japan	Chemicals and reagents
Thermo Scientific	Waltham, MA, USA	Chemicals and reagents
VWR International	Randor, PA, USA	Chemicals and reagents

3.2 Common buffers

PBS pH 7.4: made from tablets according to the manufacturer's instructions

PBST: PBS buffer with the addition of 0.05% v/v Tween 20

Tris-based buffer: 50mM Tris pH 7.4, 50mM NaCl, 5mM EDTA, 0.02% w/v sodium azide

3.3 List of computational packages

ApE	DNA sequence analysis
BLAST	Sequence alignment
CCP4[160]	X-ray crystallography data processing
Chimera[9]	Structural visualisation
Coot[161]	Structural visualisation
FIJI (ImageJ)	Image processing
GraphPad (Prism)	Statistical analysis
IMAGIC[162]	Electron microscopy image processing
Pymol[163]	Structural visualisation
Relion[122]	Electron microscopy image processing
PyRosetta[164]	Protein modelling

3.4 Protein purification

3.4.1 Purification of monomeric and polymeric M-AAT from plasma

Plasma from a wild-type (PiM) AAT individual was centrifuged at 12,000 g for 20 minutes to pellet any remaining cells and debris. The supernatant was filtered using glass fibre prefilters and cellulose acetate filters. The filtered plasma was diluted 1:1 in binding buffer (20mM Tris, 150mM NaCl, pH 7.4) before loading onto the $Alpha_1$-Antitrypsin Select column (GE Healthcare) pre-equilibrated in binding buffer. The column was washed in binding buffer until a baseline UV signal was reached. M-AAT (wild-type AAT) was eluted from the column using elution buffer (20mM Tris, 2M $MgCl_2$, pH 7.4). Protein-containing fractions, as analysed by SDS-PAGE, were pooled. The protein was dialysed into 20mM Tris, pH 8 overnight at 4°C. The dialysed sample was loaded onto a 5ml HiTrap Q HP column (GE Healthcare) pre-equilibrated in buffer A (20mM Tris, pH 8). The column was washed with 30 ml of buffer A with 20mM β-mercaptoethanol to reduce M-AAT dimer, and then buffer A alone until a UV baseline was reached. M-AAT was eluted using a 0-50% gradient of buffer B (20mM Tris, 1M NaCl, pH 8) over 80 ml. Protein-containing fractions were

analysed using SDS-PAGE and a non-denaturing gel (to identify polymers) and were pooled accordingly. Monomer was buffer exchanged into PBS with 0.02% w/v azide using a 30kDa MWCO Amicon Ultra-15 Centrifugal Filter unit (Millipore) by centrifugation at 3000 *g*. Protein concentration was estimated by the absorbance at 280nm using an E1% extinction coefficient of 5.2.

3.4.2 Production of *in vitro* polymers

All types of *in vitro* polymers were prepared from monomeric plasma M-AAT at 0.2 mg/ml. Heat-induced polymers were prepared by heating M-AAT in PBS at 55°C for 48 hours. Denaturant (urea and Guanidine-HCl) and acid polymers were incubated at 25°C or 37°C for 48 hours. Urea polymers and Guanidine-HCl polymers were prepared at a concentration of 4M urea and 3M Guanidine-HCl respectively in 40mM Tris-HCl pH8 buffer. Acid polymers were prepared using 0.1M sodium acetate pH 4.5. All polymers were purified from any remaining monomer using a 1ml HiTrap Q HP column (GE Healthcare) as in 3.4.1 but using a 0-100% gradient elution of buffer B over 20ml. Polymeric fractions were pooled and buffer exchanged into PBS pH 7.4 using centrifugation.

3.4.3 Generation of cleaved AAT

Plasma M-AAT was concentrated to 10mg/ml in 0.1M ammonium bicarbonate pH8 buffer. Endoproteinase Glu-C from *S. aureus* strain V8 was added at a ratio of 3:100 (w/w) enzyme to substrate and incubated at 37°C for 16 hours. Cleavage was stopped by addition of cOmplete Protease inhibitor and confirmed by SDS-PAGE. Cleaved AAT was purified from Glu-C using a Hi Trap Q HP column (GE Healthcare) and buffer exchanged into PBS.

3.4.4 Purification of monoclonal antibodies

The supernatant containing secreted IgG from hybridoma cells was diluted 1:1 in Pierce Protein G IgG binding buffer (Thermo Scientific) which had been equilibrated to room temperature. The diluted supernatant was loaded onto a 1ml or 5ml HiTrap Protein G HP column (GE Healthcare) pre-equilibrated in five column volumes of binding buffer. Bound material was eluted using Pierce IgG elution buffer (Thermo Scientific). Protein-containing fractions were neutralised by addition of 100 µl/ml 1M

Tris pH 9 and pooled. IgG was buffer exchanged into PBS using an Amicon Ultra-15 Centrifugal Filter unit (Millipore) with 30kDa MWCO by centrifugation at 3000 g. Sodium azide was added at 0.02% w/v and purified mAb aliquoted and snap frozen using liquid nitrogen for storage at -80°C.

3.4.5 Fab generation and purification

The fragment antigen binding (Fab) domain of IgG was produced using the Pierce Mouse IgG1 Fab and $F(ab')_2$ Preparation Kit (Thermo Scientific). An immobilised ficin column was prepared by addition of 0.75 ml of 33 % ficin slurry into a 0.8 ml spin column and centrifugation at 5000 g for 1 min. The resin was equilibrated in digestion buffer prepared for Fab generation (25mM cysteine/HCl). 0.5 ml of sample containing 4 mg of IgG was added to a Zeba Spin Desalting Column pre-equilibrated in digestion buffer and centrifuged at 1000 g for 2 min. The prepared IgG sample was added to the equilibrated immobilised ficin and incubated for 5 hours with end-over-end mixing at 37°C. The column was centrifuged at 5000 g for 1 min to separate the digest from the immobilised ficin. The ficin was washed with 0.5 ml Protein A binding buffer 3 times and flow-through pooled to a total of 2 ml. The digested antibody sample was added to a NAb Protein A column pre-equilibrated in Protein A binding buffer and incubated at 25°C with end-over-end mixing for 10 min. The column was centrifuged at 1000 g for 1 min. The flow-through contained Fab fragments. The column was washed with Protein A binding buffer twice and flow-through pooled with eluted Fab. 1 ml of Elution buffer was applied to the Protein A spin column and centrifuged; this was repeated twice. The elution fractions contained undigested IgG and the Fc fragments and each was neutralised with 100 µl of 1M Tris pH 8. The Protein A column was regenerated by washing in PBS with 0.02% w/v azide four times. Fab protein concentration was estimated by measuring the absorbance at 280nm using an extinction coefficient of 1.4.

3.5 Protein characterisation techniques

3.5.1 Polyacrylamide gel electrophoresis

3.5.1.1 SDS-PAGE

SDS-PAGE was performed using a NuPAGE 4-12% w/v acrylamide Bis-Tris gel (Life Technologies)[70]. Samples containing 2-4 µg of protein were prepared by mixing 1:4 with NuPAGE lithium dodecyl sulphate sample buffer (Life Technologies) and heated at 95°C for 5 min. The Spectra multicolor broad range prestained molecular weight ladder (Thermo Fisher scientific) was run alongside samples. Gels were run in 1X NuPAGE running buffer at a constant 170V for 50 minutes at room temperature. If samples were reduced, 50 mM DTT was added prior to heating. Protein bands were stained with Instant Blue Coomassie stain (Expedeon) with gentle rocking for 30 minutes to allow visualisation of protein bands.

3.5.1.2 Non-denaturing PAGE

Native gel electrophoresis was performed using a NativePAGE 3-12% w/v acrylamide Bis-Tris gel (Life Technologies)[70]. Samples containing 2-4 µg of protein were prepared by mixing 1:4 with a 50% v/v glycerol, bromophenol blue sample buffer. The gel was run in 1X NativePAGE running buffer at 30 mA for 35 minutes at room temperature.

3.5.2 Western blot

0.2 µg of protein was loaded per lane for electrophoresis on a NuPAGE 4-12% w/v acrylamide or NativePAGE 3-12% w/v acrylamide Bis-Tris gel (Life Technologies)[70]. The gel was electroblotted onto a PVDF membrane using the iBlot Dry Blotting System (Life Technologies) for 7 min. The membrane was soaked in PBS for 10 min and blocked with 5% w/v nonfat dry milk in PBS for 1 hour. The primary antibody for total AAT (rabbit polyclonal anti-AAT, made in house) and polymer AAT (mouse monoclonal 2C1) was diluted to 0.8 µg/ml and 0.2 µg/ml, respectively in PBS, 0.1% v/v Tween 20, 5% w/v BSA and 0.1% v/v sodium azide and incubated overnight at 4°C on a rocker. The membrane was washed six times in PBS, 0.1% v/v Tween 20 prior to incubation with the secondary antibodies. The Odyssey donkey anti-rabbit IR Dye 680RD and donkey anti-mouse IR Dye 800CW (LI-COR Biosciences) secondary

antibodies were diluted 1:10,000 in PBS, 0.1% v/v Tween 20 and 5% w/v nonfat dry milk and incubated for 1 hr, protected from light. The membrane was washed six times with PBS, 0.1% v/v Tween 20 and six times with PBS, in the dark prior to scanning using the LICOR Odyssey scanner.

3.5.3 ELISAs

For sandwich ELISAs, flat-bottomed 96 well EIA/RIA plates (Costar 3590) were coated overnight at room temperature with 50 µl per well of antigen purified rabbit polyclonal anti-antitrypsin antibody at 2 µg/ml in PBS, pH 7.4[70]. The wells were washed three times with 300 µl per well of washing buffer (0.9% w/v NaCl, 0.05% v/v Tween20) and blocked overnight at 4°C with 300 µl per well of blocking buffer (PBS, 0.25% w/v BSA, 0.05% v/v Tween20, 0.1% w/v sodium azide). Standards (heat polymer) and unknown samples were serially diluted across the plate in blocking buffer. Samples were incubated for 2 hours at room temperature and then washed three times with washing buffer. The binding antibodies 3C11 (total) or 2C1 (polymer) were diluted 1:100 or 1:50 respectively from culture media supernatant in blocking buffer and 50 µl incubated at room temperature for 2 hours. After washing, bound monoclonal antibodies were detected with rabbit anti-mouse HRP-labelled antibody diluted 1:20,000 in blocking buffer without sodium azide for 1.5 hours protected from light. The wells were washed and the reaction developed for 10 min with 50 µl TMB substrate solution. The reaction was stopped with 50 µl of 1M sulphuric acid and the absorbance at 450nm measured in a plate reader. For competition ELISAs, plates were coated with 2 µg/ml of antigen (e.g. plasma M-AAT monomer), binding antibodies were serially diluted from 2 µg/ml across the plate and HRP-labelled antibodies were added at 0.03 µg/ml.

3.5.4 Visualisation of polymers by negative stain electron microscopy

3.5.4.1 AAT polymer sample preparation for EM

AAT polymers were buffer exchanged into Tris-based buffer using an Amicon Ultra Centrifugal Filter unit (Millipore) with 30kDa MWCO by centrifugation at 10,000 g.

3.5.4.2 Labelling of AAT polymers with Fab for visualisation by NS-EM

A three times molar excess of Fab was incubated with AAT polymer for 3 hours at room temperature. The sample was dialysed at 4°C overnight into Tris-based buffer using a pre-conditioned 1ml 300 kDa MWCO Float-A-Lyzer G2 Spectra/Por dialysis device (Fisher Scientific) to remove unbound Fab. The sample was concentrated using a 10kDa concentrator to around 0.2 mg/ml.

3.5.4.3 Preparation of grids for NS-EM

The continuous carbon film of 400 mesh copper (CF400-Cu) grids (Electron Microscopy Sciences) was glow-discharged at 30 mA for 1 min using a PECLO easiGlow Glow Discharge Cleaning System (Ted Pella)[70]. 4 µl of ~0.04mg/ml sample was applied to the grid and incubated for 1 min before blotting away excess solution with filter paper. The grid was rapidly washed twice with 4 µl deionised water before negative-staining. 4 µl of 2% w/v uranyl acetate (UA), which had been centrifuged to pellet precipitate, was applied to the grid for 1 min and excess stain blotted. The staining step was repeated and the grid allowed to air dry before imaging.

3.5.4.4 NS-EM imaging

Negatively-stained grids were visualised using a Tecnai 100kV or Tecnai 120 kV field emission gun (FEG) equipped with a Gatan CCD MultiScan Camera. Samples were assessed at 44,000X or 67,000X nominal magnifications for the Tecnai 100kV or 120kV, respectively.

4.0 Generation and characterisation of monoclonal antibodies against AAT

4.1 Introduction

4.1.1 Antibodies

Immunoglobulins (Ig) play an important role in the adaptive immune system, through highly specific binding to their cognate foreign antigens and subsequent effector functions. There are five classes of immunoglobulins: IgG, IgM, IgA, IgD and IgE that differ in the type of heavy chain and their oligomeric state. IgG is the predominant class, accounting for 10-20% of serum protein[165]. Antibodies are produced by B cells, where each different B cell produces a unique antibody[166].

Antibodies consist of two identical heavy (H) and two identical light (L) chains (Figure 4.1). Each light chain consists of two globular domains, the N-terminal variable domain V_L and constant domain C_L. Similarly, each heavy chain consists of a variable domain, V_H, and three constant domains C_H1, C_H2 and C_H3. The variable domains, V_L and V_H, are involved in antigen binding and together with C_H1 and C_L form a Fab (fragment antigen-binding) arm. C_H2 and C_H3 are involved in binding immune response effector molecules and form the Fc (fragment constant)[166]. A flexible hinge in the heavy chain connects the Fab arms to the Fc region. The hinge imparts flexibility to the molecule and influences the relative conformations of the Fab arms and the Fc region[165]. All the domains are folded into a characteristic IgG fold, consisting of 2 β-sheets that are arranged into a β-barrel[166].

IgG can be further divided into four subclasses, IgG1, IgG2, IgG3 and IgG4, listed in order of decreasing abundance in serum. These subclasses are highly conserved, being more than 90% identical at the sequence level but differ in their heavy chain isotype and structure, particularly in their hinges and upper C_H2 domains[165]. These regions are involved in binding molecules of the innate immune system and therefore the different subclasses have different effector functions such as activation of complement or binding to Fc receptors[166]. The exact role of the four IgG subclasses in the immune response is not fully understood, however generally IgG1 and IgG3 are thought to be potent inducers of effector functions whereas IgG2 and IgG4 trigger more subtle responses[165].

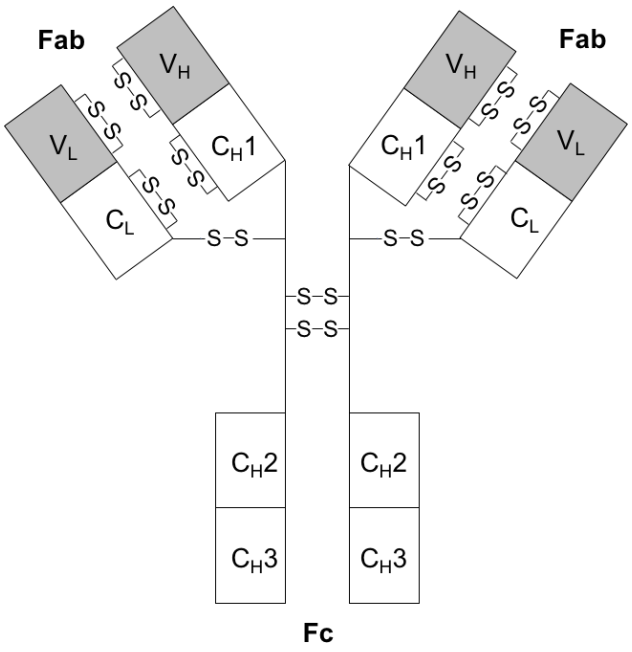


Figure 4.1. Schematic diagram of the domain structure of IgGs. *The heavy chains are composed of V_H, C_H1, C_H2 and C_H3 domains and the light chains of V_L and C_L domains. A flexible hinge links the Fab fragments to the Fc fragment.*

The variable domains are responsible for specific antigen recognition and consequently this region has the highest sequence diversity. Diversity is created by complex gene rearrangements and somatic hypermutation after exposure to antigen[167]. In particular, in each variable domain (V_L and V_H), there are three hypervariable loops, called complementary determining regions (CDRs) that make the majority of contacts with the antigen (L1, L2, L3, H1, H2, H3). H3 is the most variable of the CDR loops in sequence and length and is described as being the most important for antigen interactions[168,169]. The loops are flanked by regions of relatively constant sequence termed the framework regions (FRs).

Generally, the CDRs form the paratope (the site on the IgG at which the antigen binds) and contact around 5-8 amino acids on the protein surface. The specific region on an antigen that an antibody recognises is called the epitope. Epitopes are generally referred to as being linear/continuous or conformational/discontinuous[166]. A linear epitope exists on a single polypeptide chain and therefore persists after the protein is denatured or in smaller protein fragments. A conformational epitope is made up of amino acids brought together by 3D structure and so is only present when proteins are properly folded. A further type of epitope is a cryptic or neoepitope

in which an epitope becomes exposed or newly-formed upon a conformational change[170]. Antibodies against neoepitopes are useful in the probing of different activation states of molecules and may induce protein activation[170].

Antibodies are invaluable tools in the study of proteins in research and diagnostics. They are routinely used in laboratories through techniques such as western blot, immunohistochemistry and ELISAs. They are also critical components of diagnostic assays that detect infections or measure concentrations of biological markers, for example pregnancy tests work by detection of hCG hormone using an antibody[171]. The high specificity and selectivity of antibodies has also led to the development of monoclonal IgG-based drugs for therapeutic applications for a range of diseases, for example adalimumab is a mAb used for the treatment of autoimmune diseases[172]. As of May 2018, the FDA has approved 80 therapeutic mAbs[173]. Most of these antibodies act as therapeutics alone, whereas a few act as vehicles to deliver cytotoxic drugs to a target site[174]. The latter are called antibody-drug conjugates (ADCs) which have had particular interest in the treatment of cancer; more than 60 ADCs are in clinical trials[175]. Therefore, antibodies have useful applications in the areas of research, diagnostics and therapeutics.

4.1.2 Generating and identifying monoclonal antibodies

Serum contains many different antibodies that are specific for a variety of antigens. Antibodies that are specific for a single antigen but recognise different epitopes on the protein are called polyclonal antibodies, as they are secreted from different B-cell lineages. The mixed specificities and affinities of polyclonal antibodies made immunochemical studies more difficult. Therefore, a method to produce monoclonal antibodies (one type of antibody that originates from a single B-cell lineage and recognises a specific epitope) was developed[176]. Hybridoma cell lines provide a permanent and stable source of homogenous monoclonal antibodies (mAbs)[176,177]. The most common method of mAb generation results from the fusion of single antibody-secreting B lymphocytes (providing the antibody specificity) and B lymphocyte tumour cells (providing the immortality) (Figure 4.2)[178]. MAbs generated by this method often exhibit strong binding affinities for their antigen, removing the need for additional modifications to improve potency[177].

To generate a mAb against a specific protein, mice are immunised at regular intervals with the purified protein to stimulate an immune response. The spleen cells of the immunised mouse are then isolated and fused with cultured myeloma cells to generate hybridoma cell lines[178]. Fusion of the plasma membranes of the two cell types is generally mediated by PEG[176]. A selective medium (HAT) allows the growth of successful hybridoma fusions; it contains the inhibitor aminopterin which blocks the major nucleotide biosynthetic pathway, meaning cells must use a pathway for nucleotide synthesis that is defective in the myeloma cells but present in the mortal lymphocytes[177]. Fusion is an inefficient process: only ~1% of starting cells are fused and only 1 in 100,000 form viable hybridomas[176]. Furthermore, as hybridoma cells have two nuclei, during cell division the abnormal number of chromosomes may not be segregated equally and chromosomes may be lost. This means hybridoma cell lines can be considered unstable with loss of antibody production often resulting from loss of the chromosome carrying the functional IgG gene[176].

Antibodies are secreted into the cell media and the supernatants of the resulting hybridomas are screened using ELISA against the purified protein. Positive hybridoma clones are then redistributed to ~1 cell per well by limited dilution, iteratively cloned and re-tested by ELISA to ensure a single hybridoma clone producing a mAb with the desired characteristics has been generated[176]. The hybridoma can then be expanded as a cell line for continuous mAb production. Antibody yields after an affinity chromatography step using a Protein G column can be expected to be between 1-50mg mAb/litre of supernatant[179].

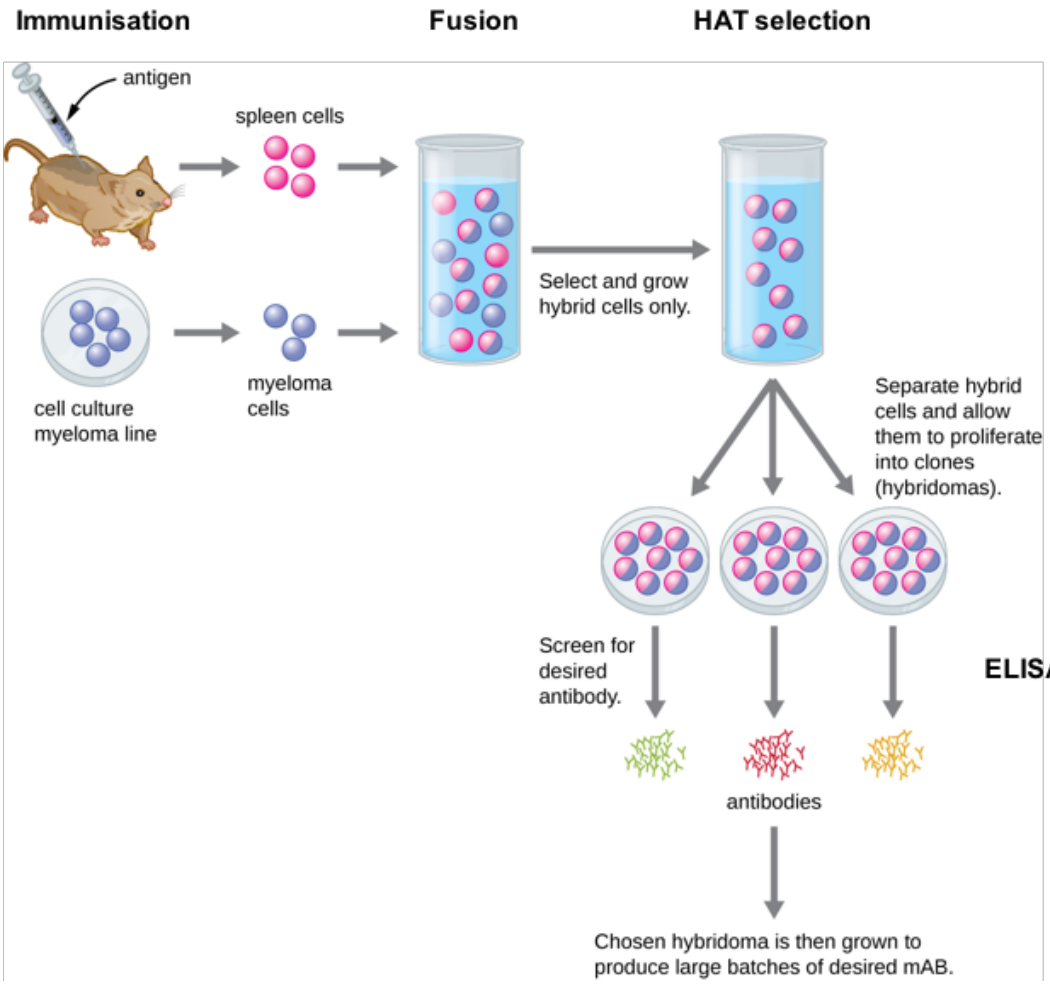


Figure 4.2. Generation of mAbs using hybridoma technology. *Mice are immunised with the antigen of interest and the spleen cells are isolated for fusion with myeloma cells. Positive fusions are selected by growth in HAT media. Antibodies are screened for the desired activity using ELISA and positive hits are cloned to ensure that only one mAb is produced. Hybridomas are then propagated for antibody production. Figure adapted from*[180].

4.1.3 Monoclonal antibodies as molecular probes of conformation

MAbs are useful molecular probes. Often, mAbs have conformational-selectivity, informing on the presence of specific epitopes but they can also exhibit functional activity, altering the behaviour of their cognate antigen[181]. They have been used to investigate a variety of proteins that undergo activation and/or conformational changes as part of their natural function or in disease states. MAbs specific for the polymeric structures of transthyretin, α-synuclein and Aβ-42 have revealed structural changes that must occur to form the aggregated state, from their binding to cryptic epitopes not present in the native monomer[182-184]. In the case of α-synuclein, the

mAbs were able to identify small, previously-uncharacterised aggregates in brain sections of patients with Parkinson's disease and Dementia with Lewy bodies, helping to inform disease pathology[183].

The conformational changes that occur within serpins in their native mechanism of protease inhibition has also been studied using mAbs. MAbs against different regions of plasminogen-activator inhibitor type-1 (PAI-1) have protected or promoted PAI-1 inhibitory activity against inactivation[185,186]. A mAb that delayed the transition to the latent conformation was found to cause structural changes in β-sheet A and the RCL, yet bound on the opposite face of the molecule[185]. This highlighted the communication between secondary elements of the serpin fold. Inactivation of anti-proteolytic activity has been achieved by mAbs that sterically block protease binding, convert PAI-1 to an inactive conformation or induce substrate-like behaviour[186-188]. One such mAb has been reported to have antithrombotic effects in the rat by binding to the bottom of the molecule and converting PAI-1 to a non-inhibitory substrate[187]. These studies have been important in informing the development of compounds targeting PAI-1 for anti-thrombosis treatment.

4.1.4 The AAT monoclonal antibody toolkit

MAb generation has been valuable in the conformational studies of AAT. Their properties have been highly informative with regards polymer structure and the polymerisation pathway but also have had wider applications such as in screening methods, detecting molecular forms *in vivo* and optimising drug preparations.

ATZ11 – specific for the Z variant of AAT

The first mAb generated against AAT was ATZ11 in 1984[189]. Mice were immunised with globular AAT inclusions (likely comprised of polymers) purified from the explanted liver of a ZZ patient. ATZ11 displayed Z specificity when plasma samples were analysed and this promoted its use in the identification of Z gene carriers, particularly heterozygotes, in an ELISA alongside isoelectric focusing. Later work using competition ELISAs showed ATZ11 interacts specifically with Z polymers and elastase-complexed AAT but lacks affinity for native, latent or RCL-cleaved AAT[190]. Accordingly, ATZ11 has a conformation-dependent neoepitope that is non-linear. Higher concentrations of antibody also resulted in the recognition of M polymers

formed *in vitro,* which indicates that it is able to bind this form, albeit with low affinity[190]. Thus, ATZ11 is specific for the polymeric and complexed form of Z-AAT.

2C1 – selective for polymers

The polymer-specific antibody 2C1 was developed by immunising mice with Z-AAT polymer made by heating purified patient-derived Z-AAT monomer[76]. 2C1 was identified to specifically recognise the polymeric form of AAT but no other conformation tested (native, trypsin-complexed, RCL-cleaved and latent) and hence was more specific than ATZ11. The polymeric epitope has recently been identified and localised to helices E and F using a protection assay with a thiol-reactive fluorescent probe[191]. This is consistent with the lack of 2C1 binding to polymer with the G117F mutation; this mutation causes a downward displacement of the F helix[192].

2C1 provided the first evidence of structural differences between AAT polymers of different origins. 2C1 binds to *in vivo* and *in vitro* heat-induced polymers but not denaturant-induced polymers. It also binds to polymers formed by different single point mutations of AAT e.g. the Z, the S, S53F Siiyama, G225R Brescia and H334D King's mutations suggesting that they share a similar structure[26,76]. However, polymers produced by the rare E75V Trento variant had a 10-fold reduced recognition by the 2C1 mAb and had a polymer size distribution more similar to denaturant polymers, suggesting different polymer forms could form *in vivo*[191].

The specificity of 2C1 has been useful in the investigation of polymer dynamics. It identified the presence of quantifiable polymers in the circulation of all ZZ-AAT homozygote patients tested[101]. They had an average concentration of 36.3 μg/ml AAT circulating polymer[101]. Plasma polymers decreased rapidly in an individual after an orthotopic liver transplantation indicating their origin from the liver. Polymer secretion from cells was demonstrated by the co-localisation of 2C1 with antibodies against proteins present in the secretory pathway in a cellular model[193]. The properties of 2C1 also make it suitable to use in the development of a diagnostic tool for the quantification and imaging of *in vivo* AAT polymers.

4B12 – a mAb with polymerisation-blocking activity

The 4B12 mAb was produced from a hybridoma screen, where the mice were immunised with monomeric Z-AAT purified from the plasma of a ZZ homozygote. It

blocks polymer formation with limited effects on the inhibitory activity of AAT against neutrophil elastase and trypsin[116,117]. The 4B12 mAb and 4B12 Fab blocked heat-induced polymerisation of AAT *in vitro* at a 1:1 molar ratio. A scFv4B12 intrabody, based on the antigen-binding variable light and heavy chains, reduced polymerisation by ~60% in a COS-7 cell model of disease as assessed by pulse-chase using radioactive ^{35}S-Met/Cys and immunoprecipitation with 2C1. It also increased the secretion of Z-AAT.

4B12 is not conformationally selective and able to bind to the native, RCL-cleaved, latent and polymeric conformations of AAT. This suggests that the epitope of 4B12 is invariant between these different conformational states. However, 4B12 does not bind to SDS-denatured material, indicating the epitope is discontinuous. Recently the epitope has been mapped using PEGylation assays and EPR to encompass residues on helix A and helix I[117]. This region is on the opposite face of the molecule to β-sheet A. As 4B12 activity was antagonised upon tethering helix I, its polymer-blocking activity is thought to arise from mAb binding causing a propagation of induced molecular changes, along the helix I-β-sheet A axis, to prematurely open β-sheet A.

4B12 has been useful in identifying a region of AAT that can allosterically regulate β-sheet A dynamics as a target for small-molecule intervention. Its antigen-altering activity has also aided in the investigation of extracellular polymer origin by preventing polymerisation of secreted monomer in cell culture medium[193].

5E3 – a mAb that recognises the polymerisation intermediate

The mAb 5E3 was identified in the same screen as 4B12 but with opposite effects on AAT[116]. 5E3 increases the rate of polymerisation of M-AAT in a dose-dependent manner[181]. It also displays conformational specificity, with a preference for polymer over monomer but no clear selectivity between denaturant and heat-induced polymers. Arginine-scanning mutagenesis identified its cryptic epitope to be in the post-helix A loop and strand 4C, which appears after monomer activation and in the resulting polymers. As 5E3 does not alter the stability of the native state, nor reduce the energetic barrier to polymerisation, it is thought to facilitate conformational conversion between the monomer and intermediate and represents a means to probe polymerisation intermediates.

1C12 – selective for the self-inserted conformation of AAT

The 1C12 mAb has been useful in investigating the presence of latent AAT in polymerisation experiments, biological samples and augmentation therapy. It displays specificity for the latent conformer (M or Z) and no detection of native, RCL-cleaved and polymeric AAT[26]. Latent AAT could be promoted *in vitro* but its kinetics and relative quantities to polymer suggested it was a by-product of polymerisation, not a precursor. Furthermore, latent AAT was not detected in a cellular model, and only little amounts were found in liver tissue samples. It was however present in augmentation therapy and the plasma of patients who had recently received this treatment. This highlights a need to optimise drug preparations to contain more inhibitory-active AAT and less inactive latent AAT.

$6C_7$-5 and 3F4 – selective for oxidised AAT

The 6C7-5 and 3F4 mAbs detect oxidised AAT (ox-AAT), but not native AAT nor RCL-cleaved/protease-complexed AAT[194,195]. The oxidation of the P1 methionine residue of AAT in the lung is thought to result from cigarette smoking, yielding a defective inhibitor[62,63]. This is thought to contribute to the progression of lung disease. The mAbs have been used to detect the presence, and measure concentations of, ox-AAT in biological samples; the concentration of ox-AAT in serum samples from patients with inflammatory diseases were higher than healthy individuals[195]. Therefore, mAbs against ox-AAT may be useful markers for evaluating oxidative stress.

5G11 – specific to peptide C-36

The C-terminal 36 residues of AAT are specifically recognised by the mAb 5G11[196]. The C-36 peptide may result from cleavage by non-target proteases *in vivo* and has been reported to stimulate pro-inflammatory activity *in vitro*[197]. 5G11 was able to detect the C-36 peptide for the first time in human lung tissue[196]. This supports the hypothesis that AAT peptides contribute to inflammation *in vivo*.

9C5 and 3C11 – non-selective mAbs

The 9C5 and 3C11 mAbs are also conformationally non-specific mAbs, binding to monomer and polymer with high affinity. 9C5 was identified in the same screen as 2C1[76]. The epitope of 9C5 is not known but it can bind to AAT at the same time as

4B12 as a HRP-linked 9C5 was used in the sandwich and competition ELISAs in the identification of 4B12[116]. The epitope of 3C11 has not been fully characterised but it competes with 4B12 for binding to AAT so is likely to have an overlapping epitope.

4.1.5 A brief overview of X-ray crystallography

4.1.5.1 Crystallisation

For X-ray diffraction experiments, protein molecules must be induced to form well-ordered crystals. This requires the sample to have high homogeneity, purity and solubility at relatively high concentrations (generally 5-10 mg/ml). The protein is brought to the point of supersaturation in a controlled manner[124]. This is done by slow equilibration of a small drop of protein (either in a sitting or hanging drop format) with a reservoir containing precipitant. Crystal formation follows two stages: nucleation where microcrystals form, and growth of a larger crystal around a nuclear seed. This process can be understood from a solubility diagram, in which a solubility curve corresponds to concentrations at which the protein solution is in equilibrium with precipitant and divides undersaturated and supersaturated states (Figure 4.3)[198,199]. The supersaturation state can be divided into 3 subzones: the precipitation zone where excess protein forms amorphous aggregates, the nucleation zone where excess protein forms crystalline aggregates and the metastable zone where crystal growth occurs. The growth of a large, well-ordered crystal should follow the red path on the solubility diagram; a small number of nuclei form in the nucleation zone and as the crystals grow, the solution returns to the metastable region for slow growth[198].

Crystallisation conditions cannot be accurately predicted for a given protein and so to increase the chances of success, multiple sample and reservoir conditions are trialled[124]. Robots such as the Mosquito have enabled nanolitre protein volumes to be accurately pipetted, making the process less tedious and allowing a wide range conditions to be trialled from limited sample volumes.

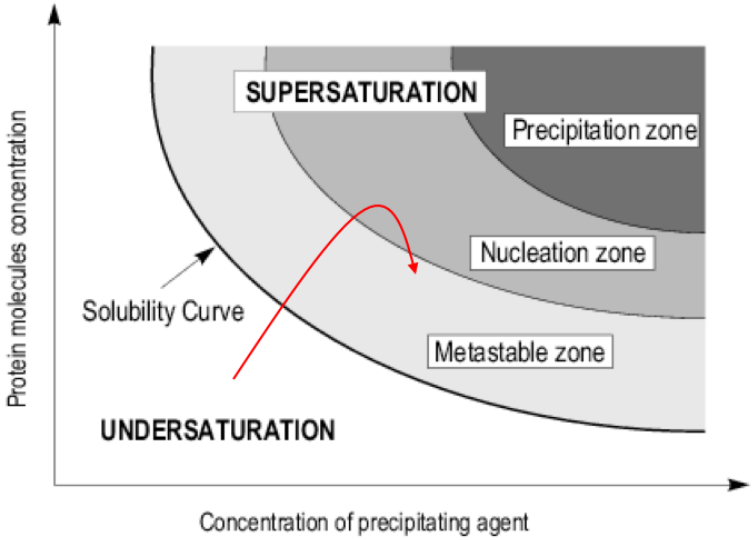


Figure 4.3. Crystallisation phase diagram. *The solubility curve divides the undersaturated and supersaturated states of protein solution, and the supersaturation state is further divided into metastable, nucleation and precipitation zones. Well-ordered proteins should follow the red arrow. Figure adapted from* [198].

A well-formed crystal lattice is a regularly repeated arrangement of protein molecules. The smallest and simplest translationally-related repeating unit of the lattice is called the unit cell. The unit cell can consist of a single protein molecule or several identical molecules related by rotational symmetry[199]. One protein molecule/complex is called the asymmetric unit. In theory, the diffraction pattern from one molecule contains all the information to generate a structure, however in practice a crystal of ordered proteins is needed to increase the SNR[124].

4.1.5.2 Data Collection

Formed crystals are soaked in cryoprotectant and flash frozen into liquid nitrogen. The frozen crystals are mounted on a nylon loop in an X-ray beam for data collection and kept cool by a stream of liquid nitrogen gas. The low temperature of the crystal reduces X-ray-induced damage, mostly caused by free radicals[200]. The most common source of X-rays used to determine structures are synchrotrons. In a synchrotron, electrons are accelerated in a circular orbit by magnetic and electric fields[124]. Intense and parallel X-rays are emitted in tangential directions to the beam as the electron path curves.

The X-ray beam irradiates the crystal and typically several hundred to several thousand diffraction patterns are collected on rotation of the crystal[199]. Diffraction patterns are most commonly recorded on CCD detectors. They comprise of a large phosphor screen that converts the diffracted X-rays into optical photons and then charge, which is the readout[124]. Depending on the crystal symmetry, a rotation less than a full 180° may be required for a complete dataset. However, the intense beams produced by third-generation synchrotrons combined with rapid high-resolution detectors mean that it is possible to collect highly redundant datasets in minutes, providing an increase in accuracy of weak reflections and the ability to account for radiation damage.

When X-rays irradiate atoms in a crystal, scattering occurs at all angles, although only certain angles (according to Braggs' law) will result in constructive interference and finite intensity[124]. These are called diffraction spots, or reflections, and form the recorded diffraction pattern. The spot positions in the diffraction pattern reflect the lattice and unit cell dimensions[124]. The dimensions are inversely related to the distance between neighbouring spots on the diffraction pattern. The intensities of each spot reflect a combination of all molecular information within the asymmetric unit.

4.1.5.3 Determination of molecular structure

As in EM, the reconstruction of a structure involves the interchange between real and reciprocal space by Fourier transformation. The crystal is an array of atoms in real space with atomic position being described by Cartesian coordinate system xyz. The diffraction pattern is recorded in Fourier space with coordinate system hkl[199]. The structure of a molecule can be determined by the summation of diffracted waves from the planes of the crystal lattice, each characterised by an amplitude and phase. Each spot in the diffraction pattern with positioning hkl has a measured intensity (amplitude is intensity$^{1/2}$) but with an unknown phase[124]. Phases are not recorded by the detector and must be determined by other methods to generate a structure.

Phases can be determined by experimental approaches or using homologous protein structures[199]. Multiple isomorphous replacement involves introducing heavy atoms, which are strong X-ray scatterers, into the crystals and then collecting diffraction patterns. The heavy atom changes the intensities of the reflections and the

differences between native and heavy reflection intensities can be used to identify the position of the heavy atoms in the crystal[199]. The amplitudes and phases of the heavy atom diffracted waves can then be interpreted and together with another different isomorphous replacement (hence the name multiple), can be used to phase the other reflections. Another experimental phasing approach is multiple anomalous dispersion. Selenium can be readily incorporated into the protein as selenomethionine and is an anomalous scatterer; it absorbs X-rays with wavelength ~0.98Å and so diffracted X-rays have a different phase from other atoms in the crystal[124]. Locating the anomalous scatterers within the unit cell can then be used to calculate phases of reflections. The most common phasing method due to the large number of structures in the PDB is molecular replacement[124]. This uses the structure of a similar protein (from which phases can be calculated) to estimate the phases of unknown structure. It is done by rotating and translating the known structure, calculating the FT of intensities (the Patterson function) and comparing to the FT of the measured intensities.

With complete amplitude and phase information of the diffracted waves, the electron density of the structure in real space can be computed. The quality of the map depends on the diffraction pattern data and the phase estimations[199]. High frequency waves are required to obtain high resolution. These are the reflection spots furthest from the centre of the diffraction pattern. The phase estimations can be improved through refining the atomic model built into the density[199].

4.1.5.4 Model building and refinement

A known sequence of amino acids can be built into the calculated electron density map using standard bond lengths and geometries as constraints. More accurate models can be built at higher resolution. At resolutions of 5Å α-helices appear as defined rods, at 4Å β-sheets can be traced, at 2-3Å the chain fold and sidechain positions can be identified and close to 1Å side chains are well-resolved and even some hydrogen atoms may be observed (Figure 4.4)[124].

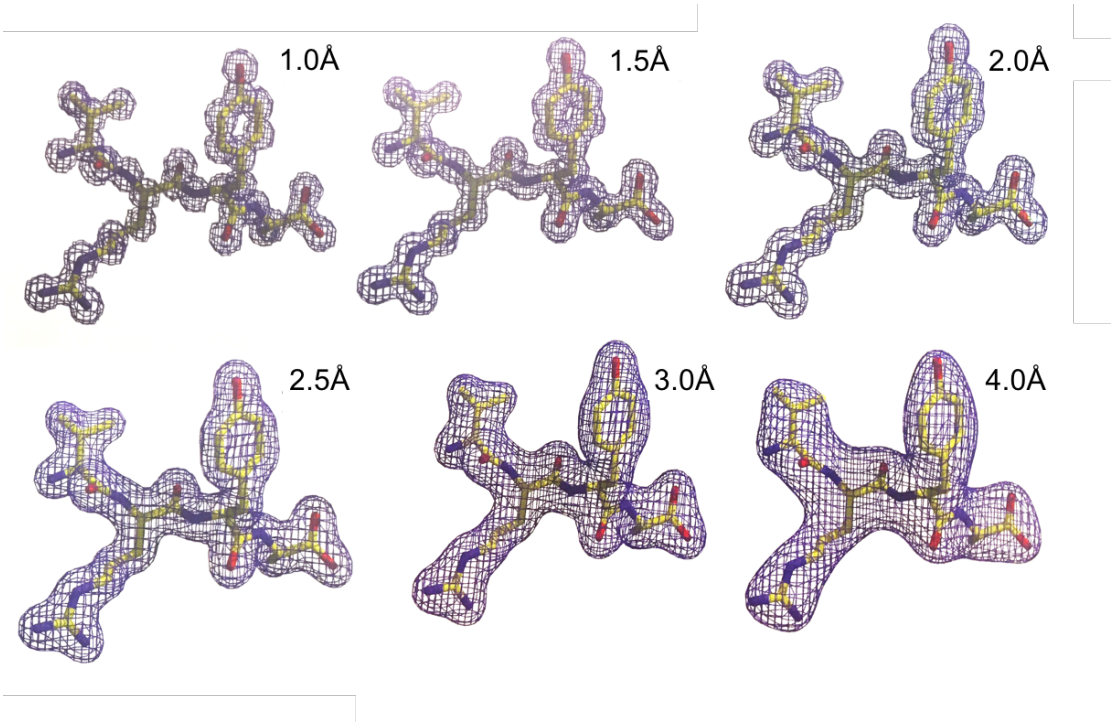


Figure 4.4. Examples of electron density at different resolutions. *At high resolution (1.0Å) atomic spheres of density can be seen and this progressively becomes more tube-like at lower resolutions. Figure adapted from*[199].

Once a model is built into the density, a theoretical set of reflections can be calculated. This can be compared to the experimental diffraction pattern. The similarity between the two is reflected in the R value (lower equals greater similarity) defined by:

$$R = \frac{\sum(|F_{obs}| - |F_{calc}|)}{\sum|F_{obs}|}$$

where F_{obs} is the observed or experimental data and F_{calc} is the calculated data[124]. The model can then be adjusted and improved to minimise the differences between the calculated and experimental diffraction patterns in an iterative refinement process. The adjustments at each refinement round change the coordinates and phases of the model to better match the data. Generally, an R value of 0.2 can be obtained from 2.5Å data[124]. To avoid over-refinement of a structure, some reflections (5-10%) are excluded from the model refinement process. The similarity of these reflections to the model are represented by an R free value which is typically 5-10% higher than the R value[124].

After refinement, structures are validated. This involves ensuring backbone dihedral angles are favoured, assessing the distribution of hydrophobic and hydrophilic amino acids and evaluating side chain rotamers. The atoms within the refined model can also be assigned B-factor values. B-factors are a measure of disorder and are obtained from analysis of the angular dependence of the diffraction pattern. Large B-factors indicate higher mobility and can reflect static disorder (structural differences between unit cells in the crystal) or dynamic disorder (molecular motion/flexibility)[199].

4.1.6 Crystallography of antibodies

The size and complexity of IgGs has limited structural analysis of the full-length proteins. Only two murine full-length IgG1 and IgG2a structures, one human full-length IgG1 and one human full-length IgG4 have been determined[201]. The flexibility of the hinge region and the resulting varied domain arrangements is thought to be the major reason for not obtaining well-diffracting crystals or clear electron density for antibodies. Accordingly, there have been structures of hinge-deleted antibodies[202,203]. Similarly, there are hundreds of structures of Fab or Fc regions, alone or in complex with binding partners that have been obtained. These have helped to determine the sequence of Fab hypervariable loops, define antigen epitopes and explain Fc effector functions.

4.1.7 Aims of this screening project

The aim of the hybridoma screening project was to produce mAbs that ideally would show novel characteristics in their recognition of AAT conformers. The intention was that these new mAbs would complement the existing mAb toolkit and provide information about polymer structure, polymer types and/or the polymerisation pathway through characterisation of their epitopes. An improved 2C1 antibody was also of interest as 2C1 has a relatively low affinity for polymer and is secreted by a low-mAb-producing hybridoma cell line. In this project, mice were immunised with either a mixture of plasma M-AAT-based acid- and urea-induced polymers or plasma polymers isolated from the plasma of donors homozygous for the Z allele. MAbs of interest were characterised immunologically, biochemically and structurally.

4.2 Methods

4.2.1 Production of the monoclonal antibodies

4.2.1.1 Immunisation

Mice immunisations and isolation of the spleens were carried out by Juan Perez (University of Malaga) as previously described[204]. A total of 12 mice were immunised, six of those with plasma ZZ-AAT polymers and six with a mixture of plasma M-based acid/urea polymers. Each mouse was immunised on five occasions spaced three weeks apart and each with 10 µg of protein. The first dose was intraperitoneal (IP) with the protein emulsified in Freund´s complete adjuvant. The following three doses were also IP but emulsified in Freund´s incomplete adjuvant and the last dose (three days before collecting the spleen cells) was intravenous in PBS. Serum samples taken seven days after each dose were analysed by ELISA for the titre and specificity of the antisera.

4.2.1.2 Fusion of myeloma and spleen cells

The fusion of myeloma and spleen cells was carried out by Juan Perez (University of Malaga) and Elena Miranda (University of Rome). A week prior to the fusion, the P3X63Ag8.653 myeloma cells were prepared by expansion into DMEM-20 (78% v/v DMEM with high glucose, 20% v/v foetal bovine serum (FBS), 1% v/v HEPES solution and 1% v/v Penicillin-Streptomycin (all components from Sigma)) plus 1mM sodium pyruvate. The myeloma cells were split every other day, including the day before fusion, to grow enough cells for the reaction. On the fusion day, myeloma cell cultures (total 200ml) were centrifuged for 5 min at 500g and the pelleted cells resuspended in 20ml of pre-warmed serum-free DMEM. The washing step was repeated twice. Thawed spleen cells were also mixed with serum-free DMEM and washed by centrifugation as above. After washing, both cell types were resuspended in 5ml and counted. Myeloma and spleen cells were mixed at a 1:3 ratio and extra DMEM added. The cell mixture was centrifuged at 500g, 5 min and the supernatant aspirated away. 1ml of prewarmed 50% w/v PEG (Sigma) was slowly (1ml/min) added dropwise to the pellet stirring gently with the pipette tip, followed by 2ml of prewarmed DMEM and a further 10ml of prewarmed DMEM. The mixture was centrifuged and the cell pellet resuspended in 50ml of HAT-20 medium

(hypoxanthine-aminopterin-thymidine (Sigma) reconstituted in DMEM and supplemented with 20% v/v FBS, 1% v/v HEPES solution, 1% v/v Penicillin-Streptomycin and 10% of fusion and cloning supplement (Roche) at room temperature. Around 125µl of cell suspension was added per well of 2 96-well flat-bottom fusion plates and the remaining 25ml of cells diluted 1:2 with HAT-20 to fill two more fusion plates. The process was repeated to complete 5 pairs of fusion plates. Cells were cultured in a humidified 37°C, 5% v/v CO_2 incubator. Every 2-3 days hybridomas were fed by replacing half of the medium with fresh medium.

4.2.1.3 Screening of supernatant media

The hybridoma clones were screened by antigen ELISA, using 50µl of supernatant medium collected from each well after 10-15 days of culturing. A mixture of AAT monomer and polymer at 4 µg/ml was used as the antigen to coat the plates. The plates were blocked as described in ELISA (3.5.3) and incubated with 50µl of hybridoma supernatant medium as the binding antibody step. The ELISA was incubated with anti-mouse IgG conjugated with HRP and developed as before. Positive wells (i.e. hybridoma colonies producing anti-AAT antibodies) were expanded by transferring them into a 24-well plate with 1ml HT medium (hypoxanthine-thymidine (Sigma) reconstituted in DMEM and supplemented with 20% v/v FBS, 1% v/v HEPES solution, 1% v/v Penicillin-Streptomycin and 10% of fusion and cloning supplement (Roche)).

4.2.1.4 Cloning

Selected clones were subcloned by limiting dilution. Briefly, each of the positive clones in the 24-well plate was progressively diluted 1:2 in column A of a 96-well plate in HT medium. Each row was then diluted 1:2 across the plate to obtain wells with one hybridoma cell. The wells were then screened again using antigen ELISA with a positive result indicating a single hybridoma clone producing anti-AAT antibodies.

4.2.1.5 Expansion

Positive single clones were expanded once confluent into 1 ml of DMEM-20 media supplemented with 10% v/v cloning supplement, in one well of a 24-well flat bottom

plate. This was done by resuspending the cells with a pipette to detach any from the surface of the plate and diluting in fresh media. Once confluent again, these cells were expanded into up to four wells in the plate (total 4 ml). These could then be expanded at a 1:5 dilution into T75 flasks (4ml + 16ml fresh DMEM-20). If the cells were expanded too early, the cells struggled to grow and to overcome this, the hybridoma cloning supplement (Roche) was added at a 1:20 dilution. From the T75 flask, cells were expanded into T175 flasks. Hybridoma cell lines were frozen as soon as possible in the expansion and the media changed to ultra-low IgG DMEM-20 media (78% v/v DMEM with high glucose, 20% v/v low immunoglobulin FBS, 1% v/v HEPES solution and 1% v/v Penicillin-Streptomycin (all components from Sigma)) for antibody production.

4.2.1.6 Production of monoclonal antibody supernatant

The mAb hybridoma cell lines were cultured in ultra-low DMEM-20, in a humidified 37°C incubator with 5% v/v CO_2. On reaching 90-100% confluency, cultures were split into new cell culture flasks with fresh media to continue the line, or left to die and lyse releasing maximum IgG into the supernatant. The supernatant was collected and 0.02% w/v sodium azide added. Monoclonal antibodies were purified as described in the method 3.4.4.

4.2.1.7 Freezing of hybridoma cells

Hybridoma cells were frozen once confluent. They were detached from the flask by washing using a pipette and the suspension centrifuged at 500g for 5 min. The supernatant was poured off and kept for IgG purification or ELISA. The cell pellet was resuspended in 2ml of freezing media (50% v/v FBS, 5% v/v DMSO, 45% v/v DMEM-20) and 1ml aliquoted into cryovials. The cryovials were placed in a freezing container (Thermo Scientific) which allowed the cells to cool down slowly and put in the freezer at -80ºC for one day and then transferred into liquid nitrogen for long term storage.

4.2.1.8 Thawing of hybridoma cells

Hybridoma cells were thawed by putting the cryovial into a 37ºC water bath. When the cells were almost thawed, the vial was moved to the hood and the cells were rapidly transferred into culture medium.

4.2.1.9 Isotyping of monoclonal antibodies

Cell supernatents were diluted 1:10 into PBS and isotyped using the IsoStrip Mouse Monoclonal Antibody Isotyping Kit (Sigma-Aldrich) by following the manufacturer's instructions.

4.2.2 Characterisation of the mAb 1D9

1D9 mAb was purified as in method 3.4.4 and plasma M-AAT was purified as in method 3.4.1.

4.2.2.1 Sequencing of the mAb 1D9

RNA was extracted from ~5 million hybridoma cells using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The Two-Step RT-PCR Kit (Takara) was used to generate DNA encoding for 1D9. The reverse transcription reaction utilised the polyT primer as a reverse primer. For the PCR step, degenerate primers based on the heavy chain IgG1 constant region and the heavy chain FR1 region[205], and the kappa chain constant region and the kappa chain FR1 region[205] were used for the heavy and light chain, respectively. The polyT reverse primer was also used in the PCR step (together with relevant forward primers) to produce DNA products that could be sequenced using a polyT primer. All PCR products were confirmed by running a 5µl sample on a 2% w/v agarose gel stained with SYBR Safe. If one DNA band was visible, amplified DNA was treated with the Enzymatic PCR Cleanup Kit (NEB). If multiple undesired bands were visible, the whole sample was run on a 2% w/v agarose gel, the DNA band excised and purified using the QIAquick Gel Extraction Kit (Qiagen). DNA sequencing was carried out by Source Bioscience (Cambridge).

4.2.2.2 Epitope mapping using CPM

25 different previously-purified single-cysteine rAAT mutants were tested for the accessibility of the engineered cysteine to conjugate with a thiol-reactive multimodal coumarin fluorophore (CPM) in the presence and absence of purified 1D9 mAb. 40µl of each mutated rAAT at a concentration of 0.2 mg/ml in PBS was added to two wells in a µclear-plate black 96 well microplate (Greiner Bio-One). 1D9 was added to a final concentration of 0.6 mg/ml (1:1 molar ratio) to one of each pair of wells. The plate was incubated at room temperature with shaking for 10 minutes to allow the complex to form. 40µl of 1:5000 CPM reagent in PBS (20nM final concentration) was rapidly added to each well and fluorescent readings recorded every 30 seconds for 30 minutes at room temperature using a SpectraMax M5 plate reader. Excitation was at 390nm and emission at 470nm. The fluorescence curves for each residue, with and without 1D9, were integrated and the difference plotted.

4.2.2.3 Thermal shift

Plasma-derived M-AAT was diluted to 0.2 mg/ml in PBS, pH 7.4 and 10 µl aliquoted into wells of a MicroAmp Optical 96-well reaction plate (Thermo Fisher Scientific). 1D9 mAb or Fab was added in 0, 0.25, 0.5, 1, 2, and 4-fold molar ratios and the complex allowed to form for 15 minutes at room temperature. 1D9 was also added to 10 µl PBS in the same ratios to obtain unfolding curves of antibody alone and as a control for buffer, 10 µl PBS was also included. 10 µl of a 1:2500 dilution of Sypro Orange (Life Technologies) was added to each well and mixed. The plate was sealed using MicroAmp Optical Adhesive Film (Thermo Fisher Scientific) and heated from 25°C to 85°C, increasing by a 0.5°C increment every 30 seconds using an Eppendorph Real4 thermal cycler. Fluorescence was recorded at 605±15 nm. The data were normalised and fit to a two-component melting equation. Satisfactory fits were obtained with the Tm of one component set to that determined for 1D9 alone.

4.2.2.4 AAT inhibitory assay

To test the effect of 1D9 on inhibitory activity of M-AAT on chymotrypsin, plasma M-AAT was diluted to a concentration of 0.1 µM in activity buffer (20mM Tris pH8, 0.1% w/v PEG 8000, 100mM NaCl, 20mM $CaCl_2$, 0.02% w/v sodium azide) and was pre-incubated with an 8-fold molar excess of 1D9 mAb for 15 minutes at room

temperature. As a control, plasma M-AAT without 1D9 was included in the assay at the same concentration. Samples were pipetted in increasing 1 µl increments from 0 µl to 10 µl in a 96 well plate, in duplicate. A constant 10 µl volume of chymotrypsin at 0.1 µM (made up in activity buffer) was added to all wells and incubated for 20 minutes with shaking. Well volumes were brought up to 20 µl with activity buffer, before addition of 180 µl chymotrypsin substrate, succinyl-Ala-Ala-Pro-Phe-p-nitroanilide (AAPF), diluted to 200 µM in activity buffer to all wells. Absorbance readings at 405nm were recorded every 30 seconds, over a 10 minute time course to monitor the reaction. The experiment was repeated with 20-fold less of the chymotrypsin (0.005 µM) for the M-AAT/1D9 sample. The data were analysed using GraphPad: Nonlinear regression was used to find the concentration of inhibitor at which there was no residual protease activity. This value was normalised such that the value obtained for M-AAT alone was equal to 1.

M-AAT inhibitory activity against human neutrophil elastase (HNE) was also assessed using the same method but with N-methyoxysuccinyl-Ala-Ala-Pro-Val-p-nitroanilide (AAPV) as the substrate and 2-fold less HNE (0.05 µM) for the M-AAT/1D9 sample.

4.2.2.5 Polymerisation assay with 1D9

Plasma M-AAT with the endogenous Cys232 was previously labelled with Alexa Fluor 488 maleimide or with Alexa Fluor 594 maleimide (Life Technologies) and labelling efficiency assessed by Dr J. Irving[74]. The two differentially labelled M-AAT were mixed in equal concentrations of label and diluted to 0.2 mg/ml in PBS. 70 µl of M-AAT mixture was added in to wells of a white 96-well semi-skirted PCR plate (Thermo Fisher Scientific) and 1D9 added in 0.5-fold, 1-fold and 2-fold molar ratios. Polymerisation was induced by heating across a temperature gradient from 50°C to 65°C using a thermal cycler. The fluorescence signal at 520 nm, 550 nm, 580 nm and 605 nm was measured over a 16 hour time course. The acceptor fluorescence (at 605 nm) was expressed as a ratio of donor fluorescence (at 520nm) and non-linear regression performed using GraphPad.

4.2.2.6 Surface Plasmon Resonance with 1D9

Binding experiments between 1D9 mAb and plasma derived M-AAT was carried out using the Biacore X100 system (GE Healthcare Life Sciences) and the Biacore X100 control software. A new CM5 sensor chip (GE Healthcare) was primed by running filtered PBS pH 7.4 over the chip before pH scouting, monitoring electrostatic interactions to find the optimal pH for 1D9 immobilisation. 1D9 was diluted to 1 µg/ml in 50mM sodium acetate buffers at pH 4, 4.5, 5 and 5 and each condition was run over the chip with a contact time of 180s. The relative response of each buffer condition was measured and surface regeneration was achieved between runs by 50mM sodium hydroxide. 50mM sodium acetate pH 4.5 was deemed to be the optimal pH for 1D9 immobilisation. The Amine Coupling Kit (GE Healthcare) was used to immobilise 1D9 on the chip by an EDC/NHS activation step, followed by 1D9 binding and an ethanolamine wash. The concentration of 1D9 was optimised to be 25 µg/ml in 50mM sodium acetate pH 4.5 with an immobilised target of 500 RU. The final response of immobilised 1D9 on the chip was 339.5 RU. A test run to determine contact time and dissociation time parameters for plasma M-AAT binding to 1D9 was run using 25 µg/ml M-AAT in PBS pH 7.4 with a contact time of 540s. This indicated binding was near-complete within 180s and dissociation was slow, requiring longer times than the 600s allowed. Regeneration of the chip was achieved to baseline response using 10mM glycine pH 2. A multicycle run with plasma M-AAT at concentrations 3.125, 6.25, 12.5 and 25 µg/ml and a regeneration step between each concentration was run with a contact time of 180s and dissociation time of 2400s. The data best fit to a heterogeneous ligand binding model and yielded affinity constants of 2 nM and 12 nM.

4.2.2.7 Peptide inhibition ELISA experiment

A sandwich ELISA was carried out as in 3.5.3 with monomeric M-AAT serially diluted across the plate. 1D9 was used as the binding antibody, which had been incubated for an hour with a 10- or 100-fold molar excess of peptide either with the RCL sequence (EAAGAMFLEAIPMS(Lys(biotin))) or the C-terminal sequence (SIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK). 1D9 alone was used as a control. Detection was carried out as in 3.5.3.

4.2.3 Production and purification of rAAT for crystallography

4.2.3.1 The plasmid encoding AAT

The plasmid used for recombinant AAT expression was based on the pQE-30 plasmid (Qiagen). The open reading frame encoding for wild-type AAT was incorporated but with a single point mutation to remove the only cysteine (C232S). The pQE-30 plasmid incorporates a His_6-tag on the N-terminus of the construct, which was removed for crystallographic studies using PCR by A. Jagger.

4.2.3.2 Transformation

1µl of pre-formed plasmid was added to 50µl of *Escherichia coli* XL1-blue supercompetent cells (Agilent Technologies) and incubated on ice for 30 min. The cells were heat-shocked for 45 seconds at 42°C and put back on ice. 200µl of SOC broth was added and the cells were incubated for 1 hour at 37°C. The cells were then plated onto LB-agar supplemented with ampicillin at 100 µg/ml for incubation overnight at 37°C.

4.2.3.3 Bacterial growth conditions

Liquid cultures of *E. coli* were grown at 37°C with shaking in LB broth. When single colonies were required, *E. coli* were plated onto 1.5% (w/v) LB-agar and incubated overnight at 37°C. Ampicillin was added at 100 µg/ml.

4.2.3.4 Expression of recombinant AAT

A total of 4L of Terrific Broth media (Formedium) supplemented with 100 µg/ml ampicillin was inoculated 1:50 with overnight culture. The cultures were grown to OD_{600nm}=0.6 at 37°C shaking before being induced by the addition of 1mM IPTG. Expression was carried out for 24 hours at 24°C with shaking. The cells were pelleted by centrifugation at 3000g for 20 min at 4°C and the pellet resuspended in PBS with lysozyme at 1 mg/ml for freezing at -80°C.

4.2.3.5 Purification of recombinant AAT without a His tag

The resuspended pellet was defrosted and DNase I and cOmplete Protease inhibitor cocktail tablets (EDTA-free) added. The cells were lysed on ice by sonication for 10 minutes. Lysates were centrifuged at 10,000g for 30 min at 4°C. The supernatant was diluted 1 in 4 in binding buffer (20mM Tris, 150mM NaCl, pH 7.4) and 0.45µm filtered prior to column loading. A 5ml pre-packed Alpha-1-Antitrypsin Select resin (GE Healthcare) column was pre-equilibrated in binding buffer and sample loaded at 2 ml/min. The column was washed with binding buffer until the UV reading reached baseline. The protein was eluted using elution buffer (20mM Tris, 150mM NaCl, 2M $MgCl_2$, pH 7.4). Fractions were analysed by SDS-PAGE, pooled and dialysed. The dialysed sample was loaded at 1 ml/min onto a 5ml Hi Trap Q HP column (GE Healthcare) pre-equilibrated in 20 mM Tris, pH 8. The column was washed until UV baseline was reached prior to protein elution using a 0-50% gradient of 20mM Tris, 1M NaCl, pH 8. Fractions were analysed by SDS-PAGE and pooled accordingly.

4.2.4 Crystallography methods

4.2.4.1 Formation and purification of rAAT/1D9 Fab complex

Purified rAAT was incubated for 2 hours at room temperature with a 2 times w/w excess of 1D9 Fab. The complex was filtered by a 0.22µm-centrifugal filter before being purified by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (GE Healthcare) that had been pre-equilibrated in 10mM Tris pH 7.4, 50mM NaCl, 0.02% w/v azide. Protein-containing peaks were run on non-denaturing PAGE to assess the presence of the complex. Relevant fractions were pooled and concentrated by centrifuging at 3,000g using a Pall Microsep Advance Centrifugal Device 30K MWCO. An extinction coefficient of 9.9 yielded a final concentration of 10.2 mg/ml.

4.2.4.2 Extra purification of 1D9 mAb and 1D9 Fab for crystallography

1D9 mAb or 1D9 Fab were diluted into 50mM sodium acetate pH 4.5 (binding buffer) and loaded onto a 5ml HiTrap SP HP cation exchange column (GE Healthcare) pre-equilibrated in binding buffer. The column was washed until UV baseline was reached prior to elution using a 0-100% gradient of 50mM sodium acetate pH 4.5,

1M NaCl. Fractions were analysed by SDS-PAGE and pooled accordingly. Proteins were concentrated by centrifugation using a Pall Microsep Advance Centrifugal Device 10K MWCO to 10.1mg/ml for 1D9 mAb and 5 mg/ml for 1D9 Fab.

4.2.4.3 Sitting drop crystallisation screens

The rAAT/1D9 Fab complex, 1D9 Fab and 1D9 mAb were subjected to crystallisation trials at their respective concentrations. Sitting drop trays were set up in a selection of sparse matrix crystallisation screens. 80 µl of each crystallisation condition well was transferred to a 96-well MRC crystallisation plate (Molecular Dimensions) using an I-PETTE liquid handling system (Fluid X). Sitting drops of 100 nl protein, 100 nl crystallisation condition were set up using the Mosquito liquid handling system (Fluid X). Plates were sealed and stored at 16°C and monitored using a Rock Imager 1000 system (Formulatrix) set to a Fibonacci imaging schedule.

4.2.4.4 Hanging drop crystallisation screens

Larger scale hanging drop crystallisation screens were set up using 24-well VDX with sealant plates and 22 mm siliconised glass coverslips (Hampton Research) to screen around positive hits. A well volume of 0.5ml was used with 0.75µl protein to 0.75µl reservoir drops. Plates were stored at 16℃.

For the 1D9 Fab: Plates were set up varying concentrations of $MgCl_2$ and PEG 3350 and pH, such as in Figure 4.5, with Fab at concentrations of 5 mg/ml or 2.5 mg/ml.

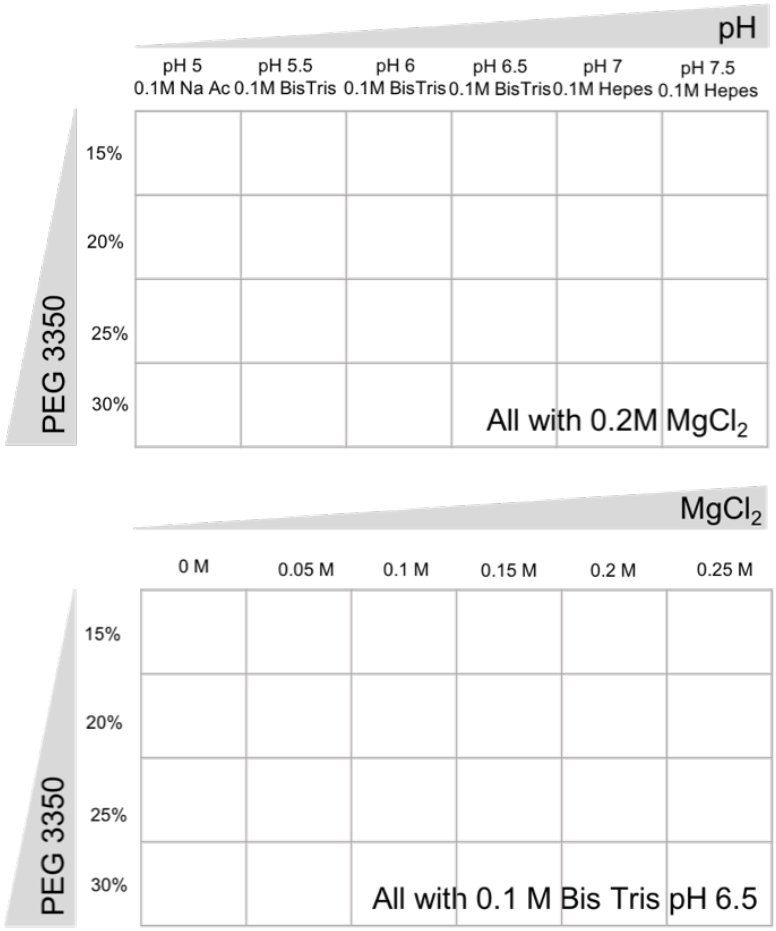


Figure 4.5. Example of hanging drop crystallisation plates set up for the 1D9 Fab. *For the rAAT/1D9 complex: Plates were set up varying concentration of PEG 4000 and the pH of sodium cacodylate such as in Figure 4.6. Finer gradations of PEG 4000 (w/v) were also trialled (10%, 11.25%, 12.5%, 13.75%, 15%, 16.25%).*

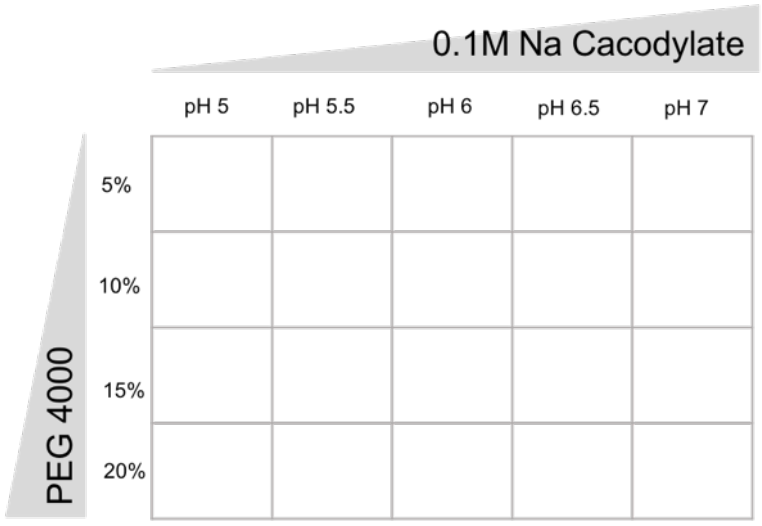


Figure 4.6. Example of hanging drop crystallisation plates set up for the rAAT/1D9 Fab complex.

4.2.4.5 Freezing of crystals

Crystals were fished from the drops using synthetic cryo-loops and dipped into cryo-protectant of mother liquor augmented with 20% w/v MPD or 20% w/v PEG 400, before being flash frozen in liquid nitrogen. Some crystals were frozen directly from the crystallisation drops.

4.2.4.6 Data collection

For the 1D9 Fab crystals: Frozen crystals were shipped to the Diamond Light Source (Oxfordshire, UK). Data were collected on the I03 beamline with a λ=1Å, beam size of 50 x 20 μm, transmission of 50% and exposure of 0.25s. 720 images were collected by rotating the crystal with a 0.5º oscillation per image, providing a redundant 360° dataset.

For the 1D9-rAAT crystals: Frozen crystals were shipped to the ESRF (Grenoble, France). Data were collected on the ID30B beamline with a λ=0.97Å, distance of 600 mm, transmission of 100% and exposure of 0.1s. 360 images were collected by rotating the crystal with a 0.5º oscillation around 180º.

4.3 Results

4.3.1 A screen of hybridomas identified five general groups of mAb reactivity profile

Mice were immunised with patient plasma Z-AAT polymers or a mixture of plasma M-AAT-based acid/urea polymers by Dr J. Perez (U Malaga, Spain). Hybridoma cell lines derived from the mouse spleen cells were first screened for AAT recognition by ELISA, in collaboration with Dr E. Miranda and Dr J. Perez. A total of 31 hybridoma cell lines secreting antibodies exhibiting this reactivity were generated. The culture media supernatant (CMS) from each cell line, which contains the secreted mAbs, was then more thoroughly screened for antigen selectivity using a sandwich ELISA. Six different antigens were used in this screen: patient plasma Z-AAT polymers, plasma Z-AAT-based heat-induced polymers, plasma M-AAT-based urea polymers, plasma M-AAT-based acid polymers, plasma Z-AAT monomer and plasma M-AAT monomer. Due to the high number of both hybridoma cell lines and antigens in the screen, end-point absorbances were measured using a single 1:200 dilution of CMS, rather than a full concentration range. This value was chosen based on previous experience with similar targets. It is important to note that the CMS was taken from hybridoma lines at different stages of culture and the absolute values of absorbance have not been normalised for antibody concentration.

The mAbs were grouped based on demonstration of a similar conformational preference profile against the six tested antigens (Figure 4.7). The antibodies belonging to Group 1 (14 mAbs) exhibited a preference for all the polymeric species tested over the monomers. The polymer-specific mAb 2C1 was included for comparison[76]. The group was subdivided into 1a (9 mAbs) and 1b (5 mAbs), where the latter had higher recognition for the monomer (based on the highest polymer signal being less than 5 times greater than the highest monomer signal). Group 2 (7 mAbs) also showed high polymer recognition but also exhibited a preference for the Z monomer over the M monomer. Group 3 (5 mAbs) had similar recognition for all the AAT conformers tested, much like the controls of 9C5, 3C11 and 4B12[76,116]. Group 4 (1 mAb) displayed a unique specificity profile with high recognition of ZZ plasma polymers, M acid polymers, Z monomer and M monomer and low recognition of Z heat polymers and M urea polymers. Group 5 (4 mAbs) had higher recognition of M acid polymers, M urea polymers and M monomer in comparison to the Z-AAT

antigens tested, suggesting that these antibodies were specific for M-AAT. The differing antigen recognition profiles of the mAbs against AAT conformers highlights the presence of structural differences between the species and the potential to gain structural or mechanistic insight through characterisation of their epitopes.

While the main goal of this part of work was to identify novel reactivities, an ancillary aim was to develop a replacement to the 2C1 polymer-specific antibody with improved characteristics, in particular a hybridoma line with increased secretion efficiency and an antibody tolerant of higher concentration. It was therefore promising to see monoclonal antibodies that were 2C1-like (Group 1). However, the hybridomas producing these antibodies failed to show the improvements desired. The most interesting generated mAb was the unique clone 1D9 that formed Group 4. This profile represents the converse of 2C1. Consequently, 1D9 was characterised in-depth using a combination of immunological, biochemical and structural techniques.

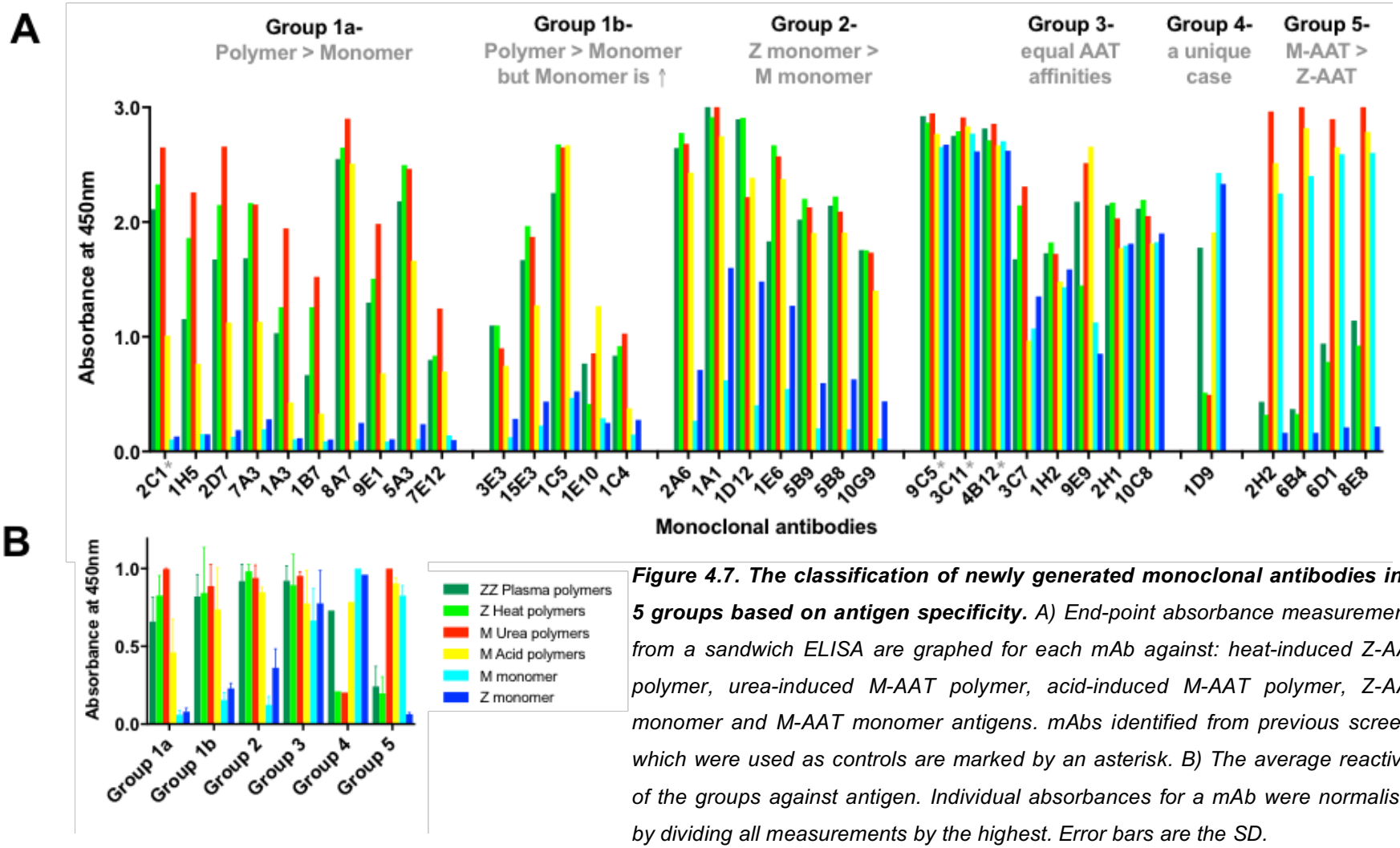


Figure 4.7. The classification of newly generated monoclonal antibodies into 5 groups based on antigen specificity. *A) End-point absorbance measurements from a sandwich ELISA are graphed for each mAb against: heat-induced Z-AAT polymer, urea-induced M-AAT polymer, acid-induced M-AAT polymer, Z-AAT monomer and M-AAT monomer antigens. mAbs identified from previous screens which were used as controls are marked by an asterisk. B) The average reactivity of the groups against antigen. Individual absorbances for a mAb were normalised by dividing all measurements by the highest. Error bars are the SD.*

4.3.2 Characterisation of 1D9 specificity using the purified protein

The monoclonal antibody 1D9 stood out in the screen for its unique combination of AAT conformer specificities. The hybridoma cell line was expanded for antibody production and 1D9 was purified from the supernatant after two weeks of secretion. 1D9 was isotyped and found to belong to the IgG1 subclass with a κ light chain. 1D9 was tested again in a sandwich ELISA with a further two antigens (plasma M-AAT cleaved between sites P4-P5 by *Staphylococcus aureus* protease GluC and liver polymers isolated from the liver of a patient homozygous for the Z allele) over a range of antigen concentrations. The running profile of these antigens on a non-denaturing gel are shown in Figure 4.8A. 1D9 was compared to the mAb 3C11 as a control, which is known to have a similar specificity for all AAT antigens. The ELISA results show that 1D9 has a 10-fold preference to monomeric AAT over polymeric forms (M monomer and Z monomer had IC_{50}s of 17±2 and 12±2 ng/ml respectively compared to IC_{50}s of 151±55, 113±20, 194±35 and 168±60 ng/ml for ZZ plasma, M-based acid, M-based urea and ZZ liver polymers respectively) (Figure 4.8B). The recognition of RCL-cleaved M-AAT was intermediate between monomer and polymer with an IC_{50} of 45±16 ng/ml. Together, these suggest that the 1D9 epitope is less available in the AAT polymer than the monomer and that part of this reduction is due to the insertion of the RCL into β-sheet A. 1D9 also displayed a surprisingly low affinity for heat-induced M-AAT polymers (IC_{50} of >400 ng/ml) compared to the other polymers. Heat-induced Z-AAT polymers and ZZ plasma/ZZ liver polymers have previously been suggested to be similar in structure[76]. These results provide the first suggestion, based on antibody reactivity, that there may be structural differences between these polymer forms. In order to interpret these differences in reactivity more fully, however, a more thorough understanding of the 1D9 epitope was required.

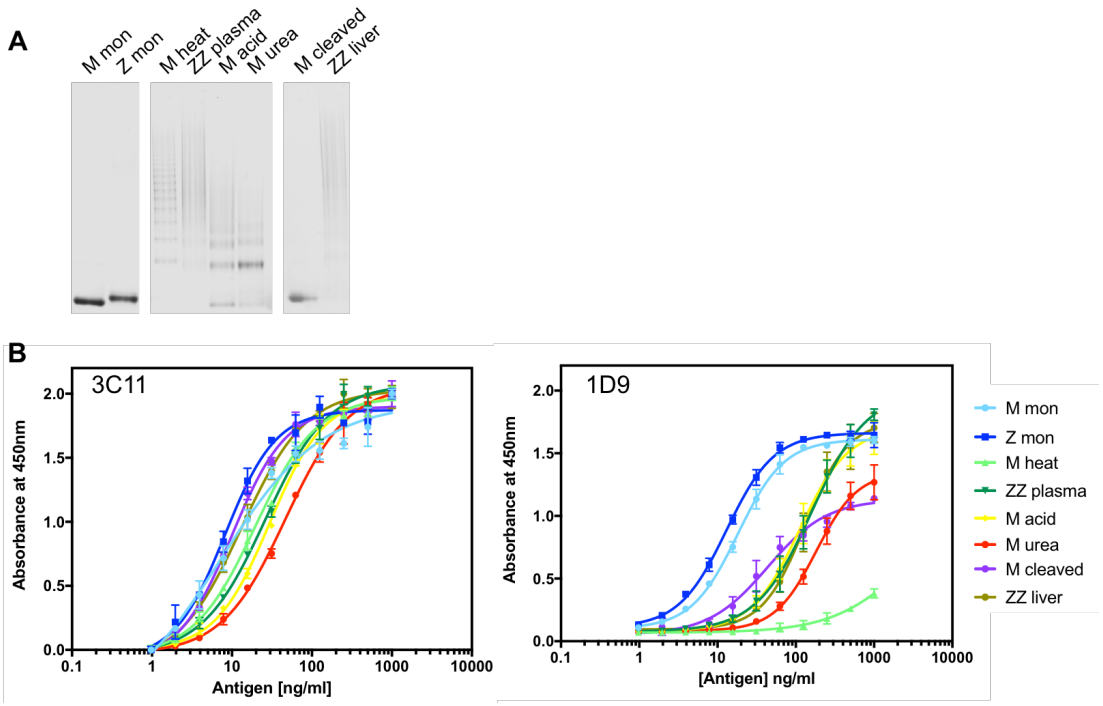


Figure 4.8. 1D9 specificity. *A) Non-denaturing gels of the eight AAT antigens screened against using sandwich ELISA. These were: M-AAT monomer, Z-AAT monomer, M-based heat-induced polymers, ZZ-AAT plasma polymers, M-based acid-induced polymers, M-based urea-induced polymers, M-AAT cleaved and ZZ-AAT liver polymers. B) Sandwich ELISA of 3C11 and 1D9 as the detection antibody against different antigens. Error bars are the standard deviation (n=3).*

4.3.3 1D9 blocks polymerisation and the inhibitory activity of AAT

As 1D9 displays selectivity for AAT monomer over polymer, it was rationalised that mAb binding could stabilise monomer-specific structural elements and thereby interfere with the process of polymerisation and/or the inhibitory mechanism of AAT.

Previously the mAbs 4B12 and 5E3 have been found to block or increase polymerisation of AAT *in vitro* respectively[116,181]. 1D9 was tested to see whether it exhibited any functional activity towards AAT. Recombinant AAT was labelled with Alexa Fluor-488 or Alexa Fluor-594, mixed in equal amounts and the increase in FRET efficiency, reporting primarily polymerisation, measured over time at different temperatures. Incubation of AAT with and without a 1:1 molar ratio of 1D9 at five different temperatures is shown in Figure 4.9. In the absence of 1D9, the FRET efficiency and the rate of change that is reporting polymerisation, increased with

higher temperatures. The greatest increase in signal was between 53.0°C and 56.8°C, indicating this temperature range encompasses the temperature required for the monomer to polymer structural transition. The formation of AAT polymers was confirmed by a fluorescent non-denaturing gel (Figure 4.9). On addition of increasing molar equivalents of 1D9, the general trend followed that the FRET efficiency and the rate of change of the FRET signal was reduced, indicating polymer formation was blocked. The greatest change in rates in the absence and presence of 1D9 was at 56.8°C, where the FRET signal and change over time was significantly reduced in the latter. At this temperature, 1D9 effectively blocked polymerisation, as demonstrated by the decrease in polymeric bands and the appearance of the 1D9-complexed monomer (marked by an asterisk) in the non-denaturing gel. The FRET signal at 60.8°C was also reduced but to a lesser extent, indicating partial 1D9 polymerisation-blocking activity at this temperature. At 63.9°C, there was little difference in signal with and without 1D9, suggesting that the higher temperature strongly drives polymerisation and/or disrupts 1D9 binding. The lowest temperature that polymerisation occurred changed from 56.8°C to 60.8°C in the absence and presence of 1D9 respectively, indicating that 1D9 is an effective inhibitor of heat-induced polymerisation, similar to the activity observed for 4B12[116].

The *in vivo* function of AAT is to inhibit proteases (predominantly neutrophil elastase) in the lungs, preventing excessive lung tissue destruction. AAT acts as a pseudo-substrate for the protease, and after RCL cleavage and subsequent insertion into β-sheet A, distorts the catalytic triad of the protease. Although not a biological target of AAT, chymotrypsin is often used in *in vitro* studies to assess inhibitory activity[64,98]. The structural rearrangement of AAT is key to its inhibitory activity[206] and if the binding of 1D9 interferes with this process then it would be expected to alter the inhibitory capacity of the protein. The inhibitory activity of AAT against chymotrypsin and human neutrophil elastase (HNE) was tested in the presence and absence of 1D9. AAT activity was measured indirectly by monitoring the rate of colorimetric substrate cleavage, which will be proportional to the amount of free protease present in solution. In the absence of 1D9, AAT was a stoichiometric inhibitor of chymotrypsin. However once bound to 1D9, AAT retained only about 4% of its inhibitory activity, with a stoichiometry of inhibition of 23±0.7:1 inhibitor (Figure 4.10). This suggests that the RCL is inaccessible for complex formation with chymotrypsin. In contrast, 1D9-bound AAT was still able to inhibit HNE relatively effectively, with a I:E of 1.8±0.04 (Figure 4.10). The extreme contrast between the SI's against HNE

and chymotrypsin is a little surprising, however the mechanism of inhibition is known to be highly sensitive to small changes, and differences in SI have been observed for the same serpin against different proteases previously[207-209].

The blocking of inhibitory activity has not been observed in any of the AAT mAbs before; although 4B12 had polymerisation-blocking activity, it only marginally increased the stoichiometry of inhibition[116,191]. This suggests the epitope includes structural elements related to the inhibitory mechanism, such as the RCL, β-sheet A and helix F[186,210,211] or the resting site of the protease[185].

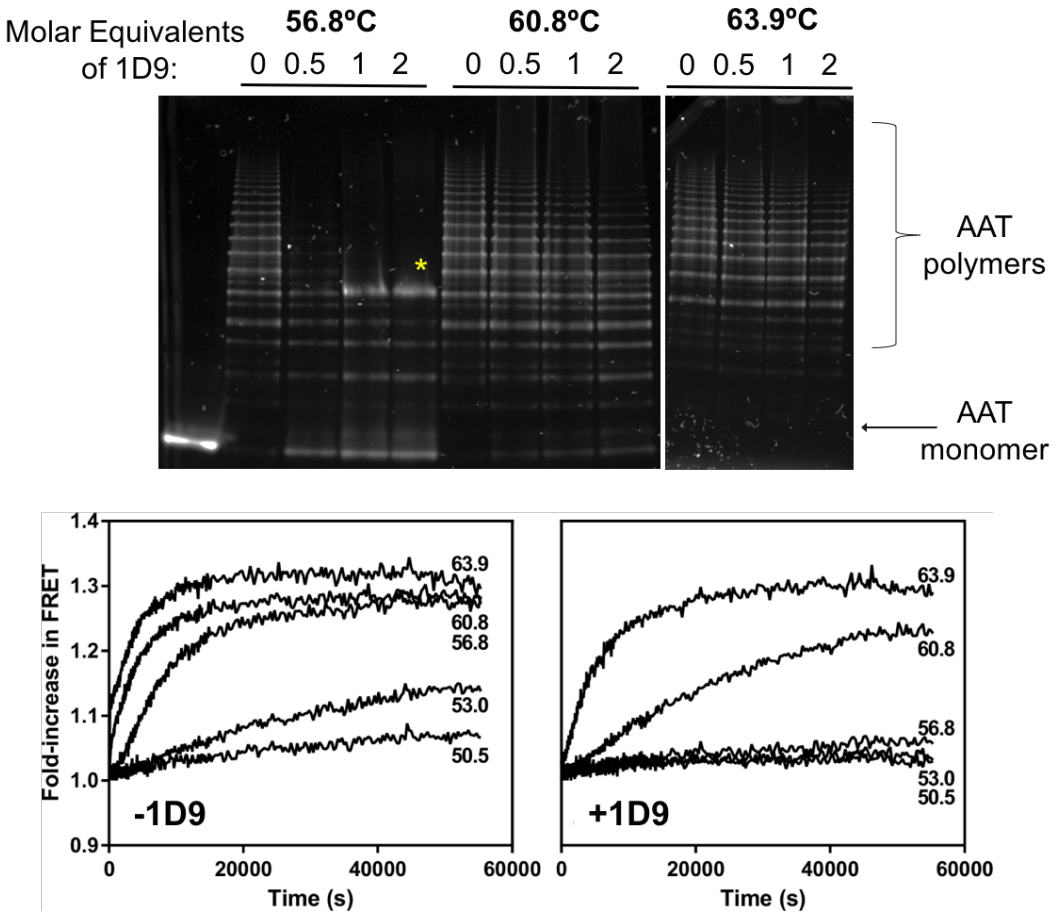


Figure 4.9. 1D9 blocks the polymerisation of AAT in vitro. *A fluorescence non-denaturing gel (top) of AAT incubated with increasing molar equivalents of 1D9 at three different temperatures. FRET analysis of fluorescently-labelled AAT molecules during heat-induced polymerisation are shown at the bottom of the figure in the absence (left) and presence of 1D9 (right). The 1D9-complexed monomer is apparent in the 56.8°C condition (marked by an asterisk). This experiment was repeated 3 times.*

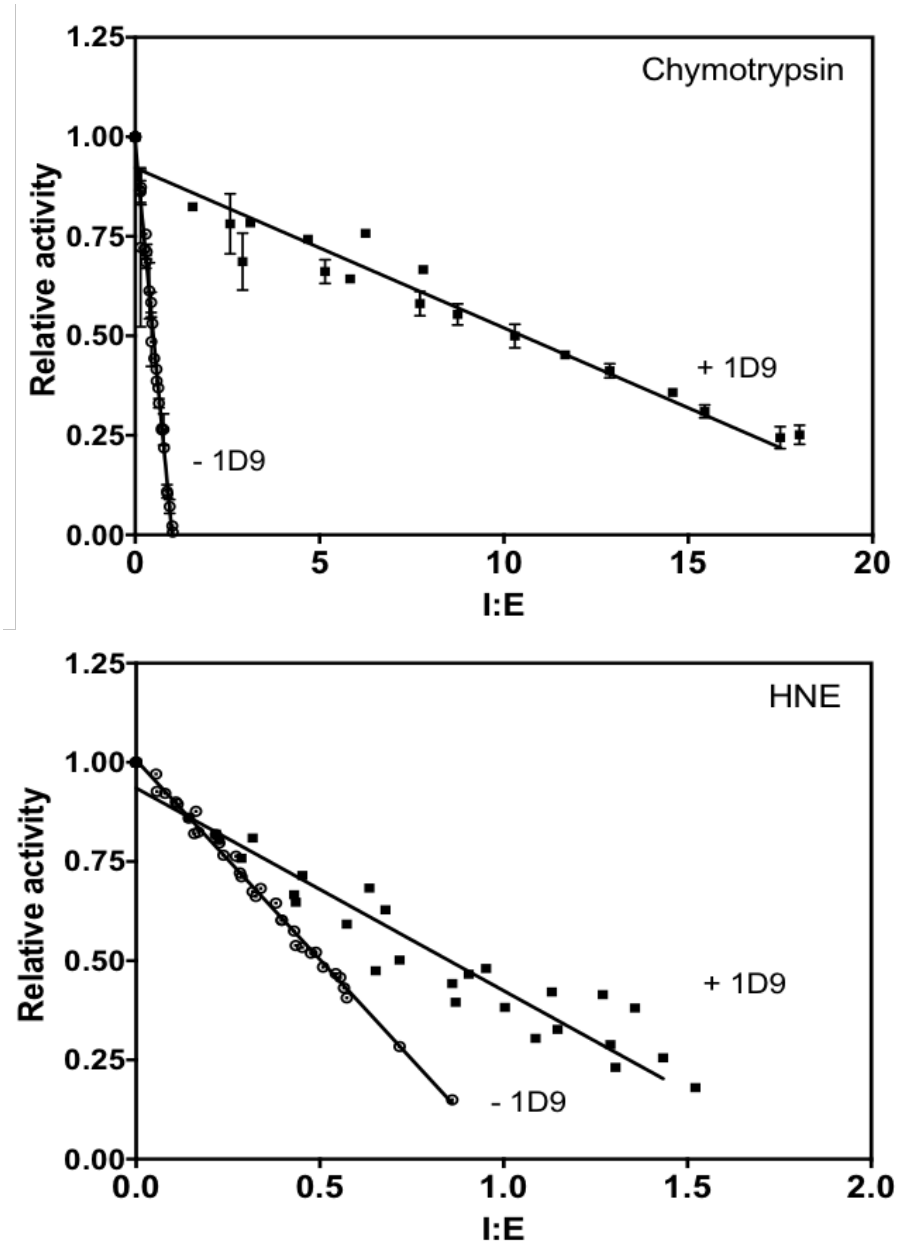


Figure 4.10. 1D9 binding alters the inhibitory activity of AAT. *Activity assays are shown of AAT against chymotrypsin (left) and human neutrophil elastase (right) in the absence and presence of 1D9. The graphs represent three experimental repeats. To account for the presence of a proportion of inactive protease molecules, the X-axis has been normalised such that in the absence of 1D9 the intercept occurs at a well-established ratio of 1:1.*

4.3.4 1D9 increases the thermal stability of AAT

The transition of native AAT to the polymerisation intermediate can be analysed in a thermal shift experiment[74]. Monomeric AAT was incubated over a temperature gradient from 25°C to 95°C in the presence of the environment-sensitive dye SYPRO Orange. SYPRO Orange undergoes an increase in fluorescence upon binding to hydrophobic regions in partially unfolded (intermediate-like) or denatured proteins. SYPRO Orange fluorescence was monitored as a function of temperature for AAT unfolding, in the absence and presence of 1D9 (Figure 4.11). In the absence of 1D9, M-AAT had a T_m of 60.7±0.4°C. The unfolding of 1D9 mAb and Fab was also measured, yielding T_m values of 70.0±0.2°C and 71.2±0.4°C, respectively. The fluorescence signal in the presence of 1D9 was deconvoluted using a two-state unfolding equation for the two-protein system to determine the midpoint of denaturation (T_m) of AAT, which can report on native state stability. In the presence of increasing molar equivalents of 1D9 mAb, the T_m of AAT increased up to a maximum of 64.8±0.6°C (SD, n=6) for four molar equivalents. Similarly, in the presence of four molar equivalents of 1D9 Fab, the T_m of AAT increased to 64.8±0.5°C (SD, n=4). The similarity of these values, despite the difference in the appearance of the curves of the complexes, is most likely due to the greater contribution to the fluorescence signal of the 150kDa mAb with respect to the 50kDa Fab. Together, these results indicate that 1D9 stabilises AAT from progressing to the intermediate state.

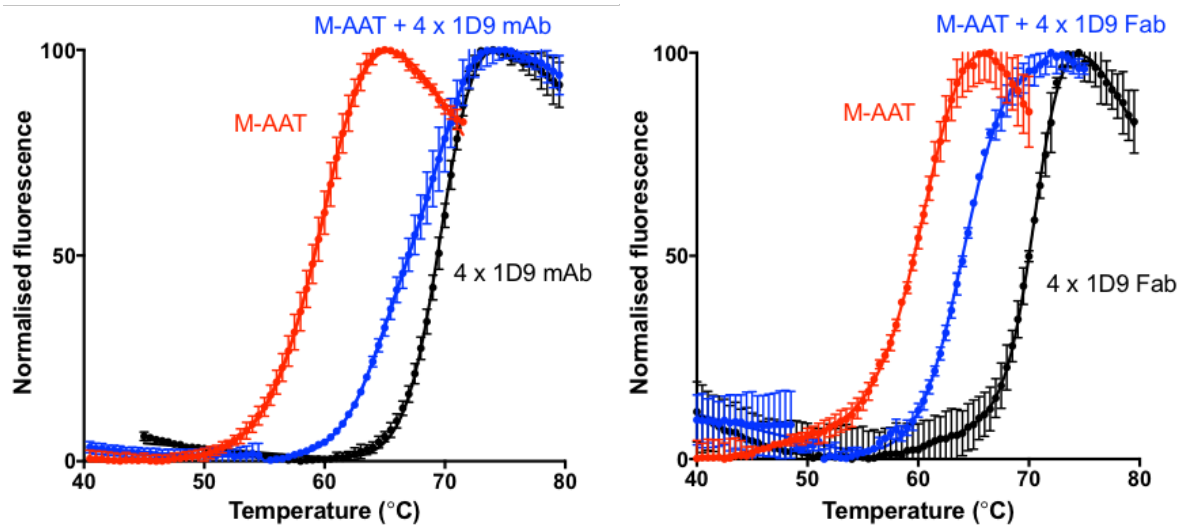


Figure 4.11. 1D9 binding stabilises the native state of AAT. *Thermal shift experiments using SYPRO Orange fluorescence to measure the melting curves of M-AAT, M-AAT with 4-fold molar excess of 1D9 mAb and 1D9 Fab. Binding of AAT by 1D9 mAb (left) and 1D9 Fab (right) were measured. Error bars denote the SD, n=6 for the 1D9 mAb experiments and n=4 for the 1D9 Fab experiments.*

4.3.5 1D9 binds near the RCL region of AAT

The effects on inhibition and polymerisation suggest 1D9 inters with conformational change of the AAT molecule. To investigate how 1D9 affects the behaviour of AAT, epitope mapping of 1D9 was carried out. First, a competition ELISA between 1D9 and 4B12 was used to see whether these antibodies had overlapping epitopes, as both mAbs have polymerisation-blocking activity. M-AAT monomer was used as the antigen, the antibody to be tested for competition was serially diluted and 4B12-HRP added at constant concentration was used for the detection system. 4B12 is known to have residues on helix A and helix I in its epitope from PEGylation and EPR experiments[116,191]. Unlabelled 4B12 was included as a positive control for competition against itself. 9C5 and 3C11 were also included for comparison, where 3C11 has previously been shown to compete with 4B12 and therefore indicates that 4B12 and 3C11 have overlapping epitopes. 1D9 did not compete with 4B12 binding, similar to that seen with 9C5 (Figure 4.12). Therefore, 1D9 binds to AAT at a different site from 4B12 and the epitope is unlikely to be around the A and I helices.

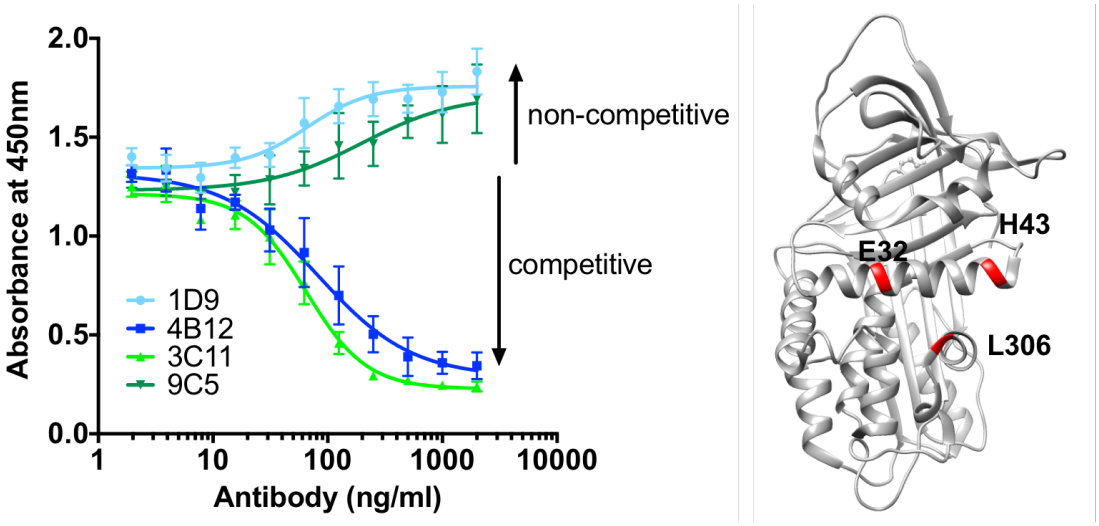


Figure 4.12. 1D9 has a different epitope to 4B12. *A competition ELISA between mAbs and 4B12-HRP shows that 1D9 and 4B12 do not compete for AAT binding. The residues coloured in red on the AAT structure (PDB ID: 1QLP[8]) indicate the location of the 4B12 epitope. Error bars represent SD for n=2.*

As 1D9 also abrogated inhibitory activity of AAT against chymotrypsin, this suggested that 1D9 might bind near to the β-sheet A or the RCL of AAT. To localise the 1D9 epitope, the same protection assay using the thiol-reactive fluorophore 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumain (CPM) that had localised the 2C1 epitope, and confirmed the 4B12 epitope, was used[191]. This molecule becomes fluorescent upon reaction of the maleimidyl group with a free cysteinyl thiol. A range of single-cysteine recombinant mutants covering the AAT molecule were tested. Each mutant was incubated with or without 1D9 antibody and the fluorescence emission on conjugation of CPM to the cysteine residue followed over 30 minutes. The difference between the integrals of the two fluorescence curves (± 1D9) was plotted (Figure 4.13). The majority of residues tested showed negligible effect of 1D9 binding on fluorescence intensity (curves for residue 43 on helix A are shown). However, residues 283 and 360, in s2C and the RCL respectively, showed a much slower increase in fluorescence intensity in the presence of 1D9. This indicates a reduced degree of fluorophore conjugation due to the interference by 1D9 and thus that they constitute part of the 1D9 epitope. Residue 346 at the N-terminal end of the RCL displayed the opposite effect; fluorophore conjugation appeared to increase in

the presence of 1D9, potentially suggestive of 1D9 binding making 346 more exposed.

The 1D9 epitope was further probed by western blot and ELISAs. A denaturing western blot showed that 1D9 recognised all AAT antigens tested, at a similar intensity to a rabbit total polyclonal antibody, except for cleaved AAT (Figure 4.14A). Binding of an epitope after denaturation is suggestive of a predominantly linear, rather than conformational epitope. The lack of recognition of cleaved AAT by a denaturing western blot suggests that the 1D9 epitope is at least partly encapsulated by the 40 residues lost after GluC-cleavage of AAT at position 354. This region would be separated from the molecule by SDS-PAGE but would still be present under native conditions, explaining the recognition of cleaved AAT by ELISA (Figure 4.8) and a non-denaturing western blot (not shown). This finding is in keeping with the identification of residue 360 being near or within the epitope.

To identify whether 1D9 bound to the C-terminal fragment of AAT or more specifically the RCL, ELISAs were carried out against M-AAT monomer after pre-incubation of 1D9 with a 10-fold or 100-fold molar excess of a peptide corresponding to the RCL or corresponding to the C-terminal 35 residues. The results showed that neither peptide affected the recognition of monomer by 1D9 (Figure 4.14B). Therefore, it is unlikely that the 1D9 epitope is solely linear.

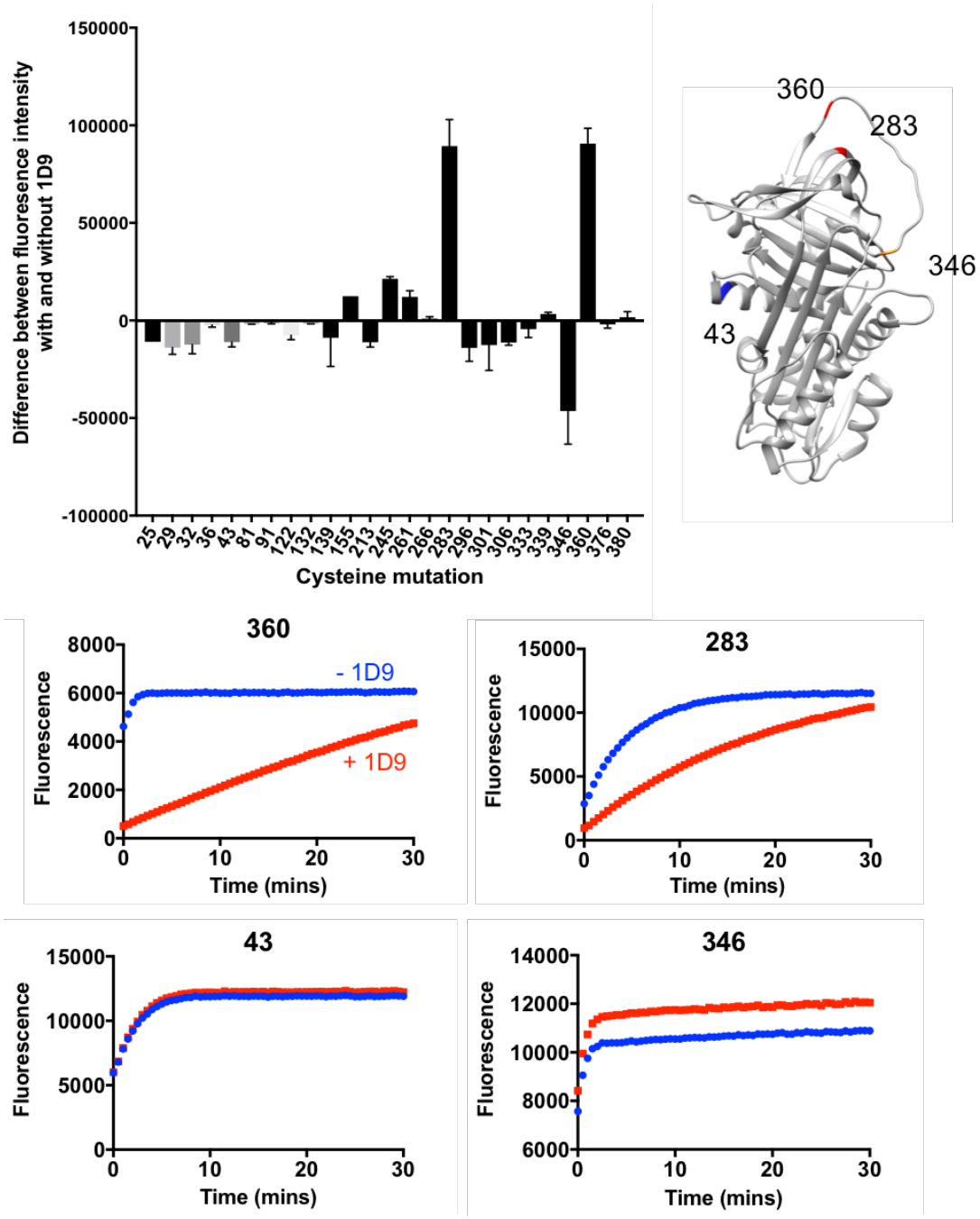


Figure 4.13. Localisation of the 1D9 epitope. *The fluorescence of the CPM dye was monitored over time as it bound to single cysteine mutations in AAT, in the presence and absence of 1D9 mAb. The difference between the integrals of the fluorescence curves was plotted. The cartoon structure of AAT highlights the position of positive epitope hits (283 and 360), as well as 346 and negative control 43. Error bars represent SD for n=2. The individual fluorescence curves for 4 residues are shown below in the absence (blue) and presence (red) of 1D9.*

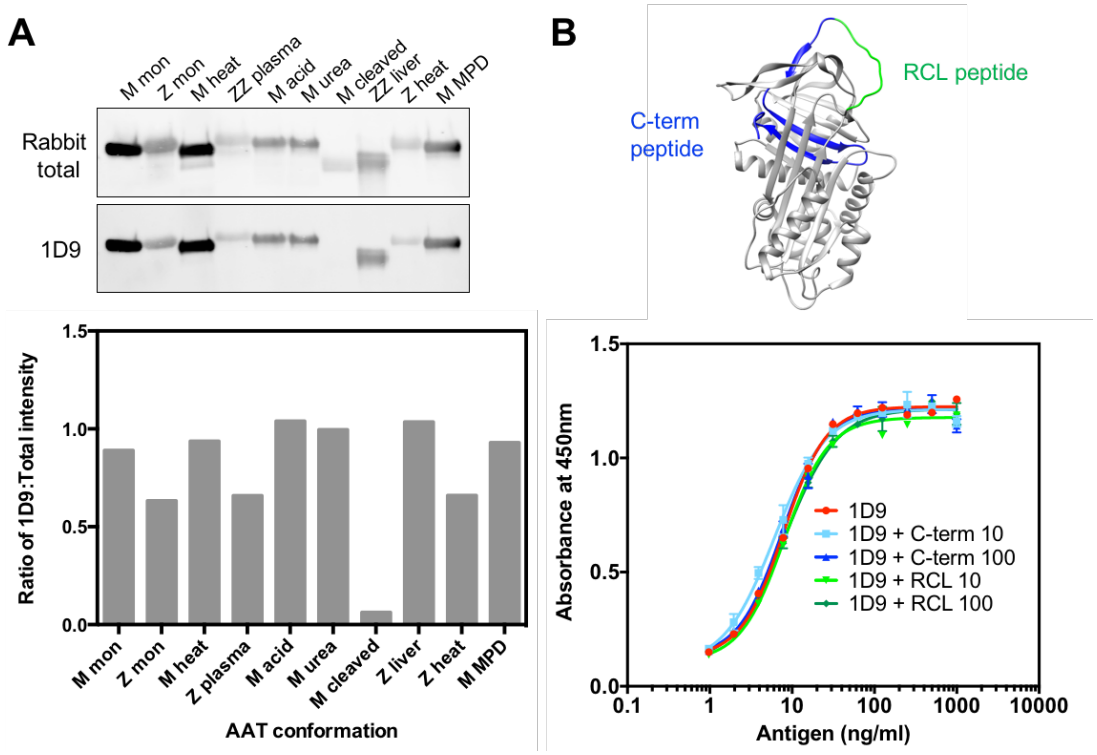


Figure 4.14. Further investigation of the 1D9 epitope. *A) A denaturing western blot of 10 AAT antigens (M-AAT monomer, Z-AAT monomer, M-based heat-induced polymers, ZZ-AAT plasma polymers, M-based acid-induced polymers, M-based urea-induced polymers, M-AAT cleaved, ZZ-AAT liver polymer, Z-based heat-induced polymers and M-based MPD-induced polymers) detected with a rabbit polyclonal antibody and 1D9 monoclonal antibody. The ratios of the band intensity for each antigen as detected by 1D9 and the rabbit polyclonal antibody were calculated using Fiji and were plotted. RCL-cleaved M-AAT was not recognised by 1D9. B) A competition ELISA using 1D9 mAb pre-incubated with peptides corresponding to the C-terminal residues S359-K394 (blue) or RCL residues E346-S359 (green) of AAT. Both the C-terminal and RCL peptide did not interfere with 1D9 binding to M-AAT.*

4.3.6 1D9 has nanomolar affinity for monomeric AAT

Surface plasmon resonance (SPR) is a powerful technique to detect binding events and was used to characterise the binding of 1D9 to monomeric M-AAT. In SPR, polarised light, shone at a certain angle, undergoes total internal reflection through a glass prism[124]. The backside of the glass prism is covered with a gold film, together forming the sensor chip. Light is absorbed by electrons in the gold film causing them to resonate (known as surface plasmons). This results in an intensity loss in the

reflected beam at the detector. In an SPR experiment, a ligand is attached to the gold side of the sensor chip, whilst the analyte molecule is free in solution and is passed over the sensor. As molecules bind to ligand, the refractive index close to the chip changes. This in turn alters the angle of minimum reflective intensity in a manner which is proportional to the mass bound, allowing binding events to be detected in real time.

As antibodies are routinely attached to sensor chips in SPR experiments, 1D9 mAb rather than AAT was immobilised via amine coupling. Firstly, the pH at which to dilute 1D9 for immobilisation was investigated (Figure 4.15A). This was done by measuring the increase in response units (RUs, reporting a change in the refractive index) solely from electrostatic interactions. It is desirable to have the highest interaction with the sensor at the most neutral pH. At pH 5.5, 1D9 did not alter the surface properties of the chip, indicating at this pH, the antibody was not sufficiently positively charged to interact with the negatively-charged carboxylate groups on the sensor chip surface. At pH's 5.0, 4.5, 4.0 there was an interaction with the chip surface indicated by the increase in RU. Therefore, at these pHs, 1D9 had a high enough positive-charge to promote interactions with the chip. As the interaction was similar between pH 4.5 and 4.0, and greater than at pH 5.0, pH 4.5 was selected for the immobilisation protocol. Immobilisation uses a 3-step procedure involving EDC/NHS, 1D9 and ethanolamine washes (Figure 4.15B). The EDC/NHS activates the chip by modifying the carboxymethyl groups of the dextran to N-hydroxysuccinimide esters. 1D9 is then flowed over the chip in small injections and covalently attaches to the esters via amine groups. The baseline RU between injections is measured until a value near the specified signal is met (in this case it was 500RU). The ethanolamine wash deactivates excess carboxylate reactive groups on the chip (as it is also a primary amine) and dissociates any remaining electrostatically bound 1D9 molecules to give the 1D9 immobilised on the chip (339RU).

Monomeric M-AAT at 25 µg/ml was flowed over the chip with the immobilised 1D9 to determine estimations for the contact time (association) and PBS wash time (dissociation) required for the interaction (Figure 4.15C). From this test run, it was apparent that the association rate was very quick, reaching the maximal RU after around 180 seconds. This signal was around 125RU indicating that not all the immobilised 1D9 was able to bind monomer. The dissociation time was much slower

with the RU reducing in value by only 35% in 600 seconds. Together, these association and dissociation times indicated 1D9 had a high affinity for AAT.

A binding experiment with a range of AAT concentrations was carried out with a 180 second contact time and 2400 second dissociation time, based on the times discovered in the test run. These binding curves reproduced what was previously seen: a fast association and slower dissociation (Figure 4.15D). The sensorgrams showed significant curvature, meaning the rates of binding and dissociation were measurable. Therefore, kinetic analysis was carried out using the Biacore software using various binding models. The simplest model (1:1 binding) was tested first with progressively more complex binding models. The heterogeneous ligand model gave the best visual fit across all concentrations with the smallest residuals. This model represents independent binding of one analyte species to two different ligand sites and is in keeping with an antibody and its two antigen-binding domains if the two binding events are treated as independent. This would be the case if the cross-linking reaction resulted in changes to the binding site of a sub-population of the molecules. The analysis yielded one binding site with $ka_1 = 6.0x10^4 M^{-1}s^{-1}$ and $kd_1 = 7.0x10^{-4} s^{-1}$ and the second binding site with $ka_2 = 4.6x10^5 M^{-1}s^{-1}$ and $kd_2 = 8.6x10^{-5} s^{-1}$. This gives binding constants of $K_{D1} = 1.9nM$ and $K_{D2} = 11.6nM$, respectively. Similar results were observed when the binding experiment was carried out with a fresh sensor chip. Affinity analysis was not carried out with these data as the sensorgrams did not reach a steady state. This is an alternate method for deriving an affinity constant by measuring the response of an interaction that has reached equilibrium at differing analyte concentrations. In theory, the same K_D values would be obtained.

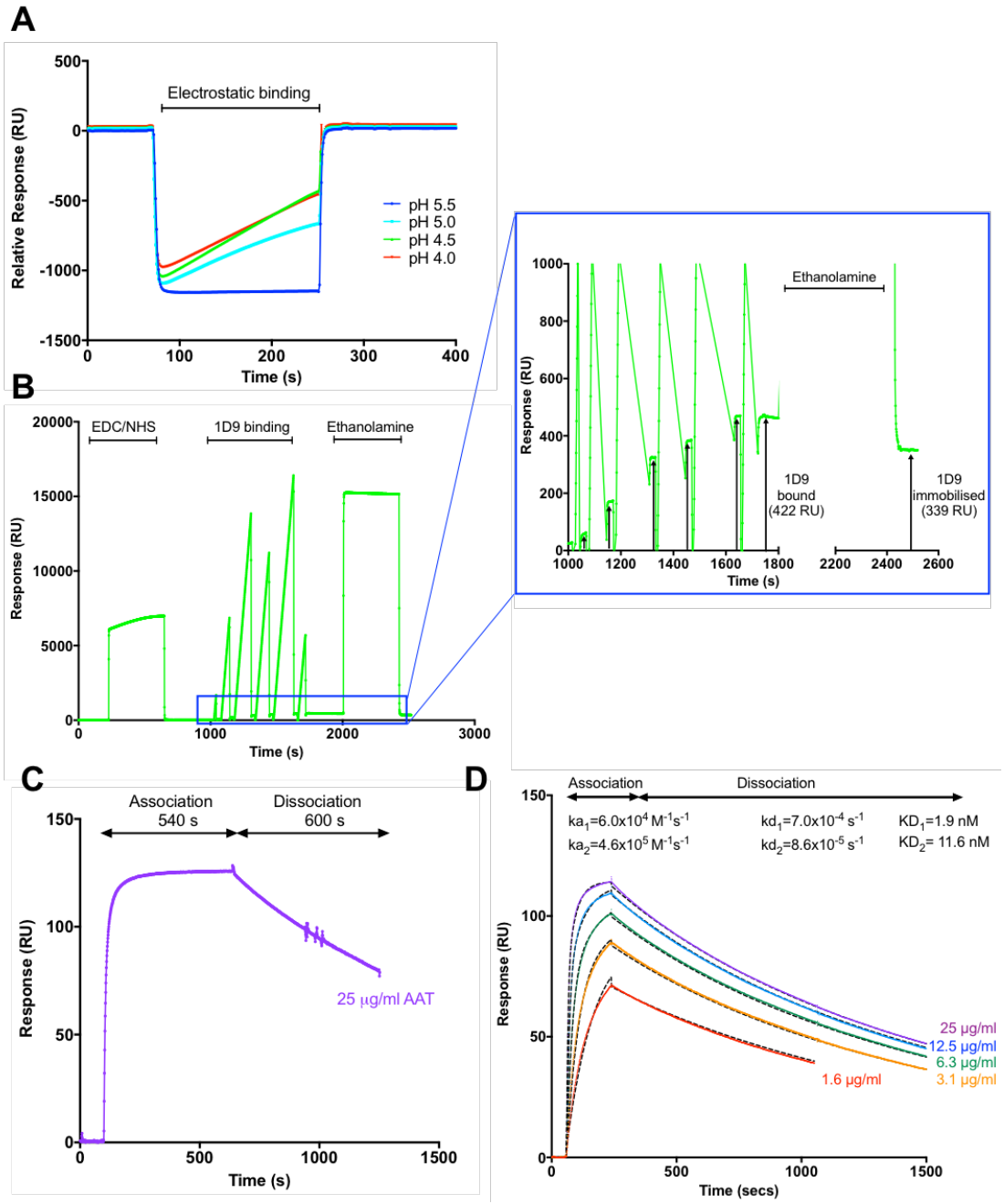


Figure 4.15. Characterisation of the binding between monomeric AAT and 1D9 using SPR. *A) A sensorgram measuring electrostatic binding of 1D9 diluted in 50mM sodium acetate at different pHs to the sensor chip. B) A sensorgram displaying the three-stage immobilisation protocol resulting in 339RU of 1D9 immobilised. C) A sensorgram of AAT binding to immobilised 1D9 with 540s association time and 600s dissociation time. D) Kinetic analysis using a heterogeneous ligand model of a binding experiment with AAT at a range of concentrations.*

4.3.7 Preparation of samples for crystallography studies

A structural understanding of the interaction between 1D9 and monomeric AAT would be highly informative in elucidating how 1D9 exerts its functional effects on AAT. To this end, the complex between recombinant AAT and the Fab fragment of 1D9 was made and purified for crystal trials.

Native AAT is glycosylated at three sites and has a single Cys residue at Cys232 that can form disulphide-linked dimers. Both these properties are not conducive for crystallisation. In addition, the typical recombinant protein used in the laboratory has an N-terminal hexahistidine tag for facilitating purification; there is a chance that this tag may interfere with crystallisation. For these reasons, a recombinant Cys232Ser AAT variant without a hexahistidine tag was expressed in *E. coli*. C232S AAT was purified using a two-step purification protocol with the $Alpha_1$-Antitryspin select column followed by a Q-sepharose column (Figure 4.16A). The 1D9 mAb was purified from the culture media of the 1D9 hybridomas using a Protein G affinity column (Figure 4.16B). Pure 1D9 mAb was digested by Ficin into Fab and Fc fragments, and then the Fc component was removed by passing the sample through a Protein A column. Fab fragments form a 1:1 complex with rAAT and as they lack the hinge regions and divalent binding of the source antibody, they provide a better-defined, less flexible complex for crystallography. The rAAT-1D9 Fab complex was formed by mixing an excess of Fab fragment with rAAT and incubating for two hours. The complex was then purified using size-exclusion chromatography (Figure 4.16D). A non-denaturing gel was used to assess complex formation. Eluted fractions showed the presence of 1D9 Fab-rAAT complex together with either the presence of some free rAAT or partial dissociation of the complex on the gel. The fractions containing the highest proportion of complex were concentrated to 10.2 mg/ml for crystallography trials.

The 1D9 mAb and 1D9 Fab were also subjected to crystallography trials as a 1D9 structure could be used to model the rAAT interaction if the complex did not produce usable crystals. In test runs of an extra purification step of 1D9 Fab using size-exclusion chromatography, it was apparent that the Fab was interacting with the column beads and this affected the separation (not shown). Furthermore, the 1D9 Fab did not appear to interact with the Q-sepharose column, unlike the 4B12 Fab, at pH values of 8 and 8.7 (not shown). Therefore, an S-sepharose column was used to

purify the 1D9 mAb and 1D9 Fab by cation-exchange chromatography (Figure 4.16C). Antibodies are generally regarded as stable at low pH for sufficiently short periods of time and this was confirmed from the regeneration step in the SPR experiments described above. Smaller molecular weight contaminants could still be observed by SDS-PAGE after elution, particularly in the 1D9 Fab sample. As no such products were observed during size-exclusion chromatography of the complex, they may be secondary digestion products that did not result in dissociation of the molecule.

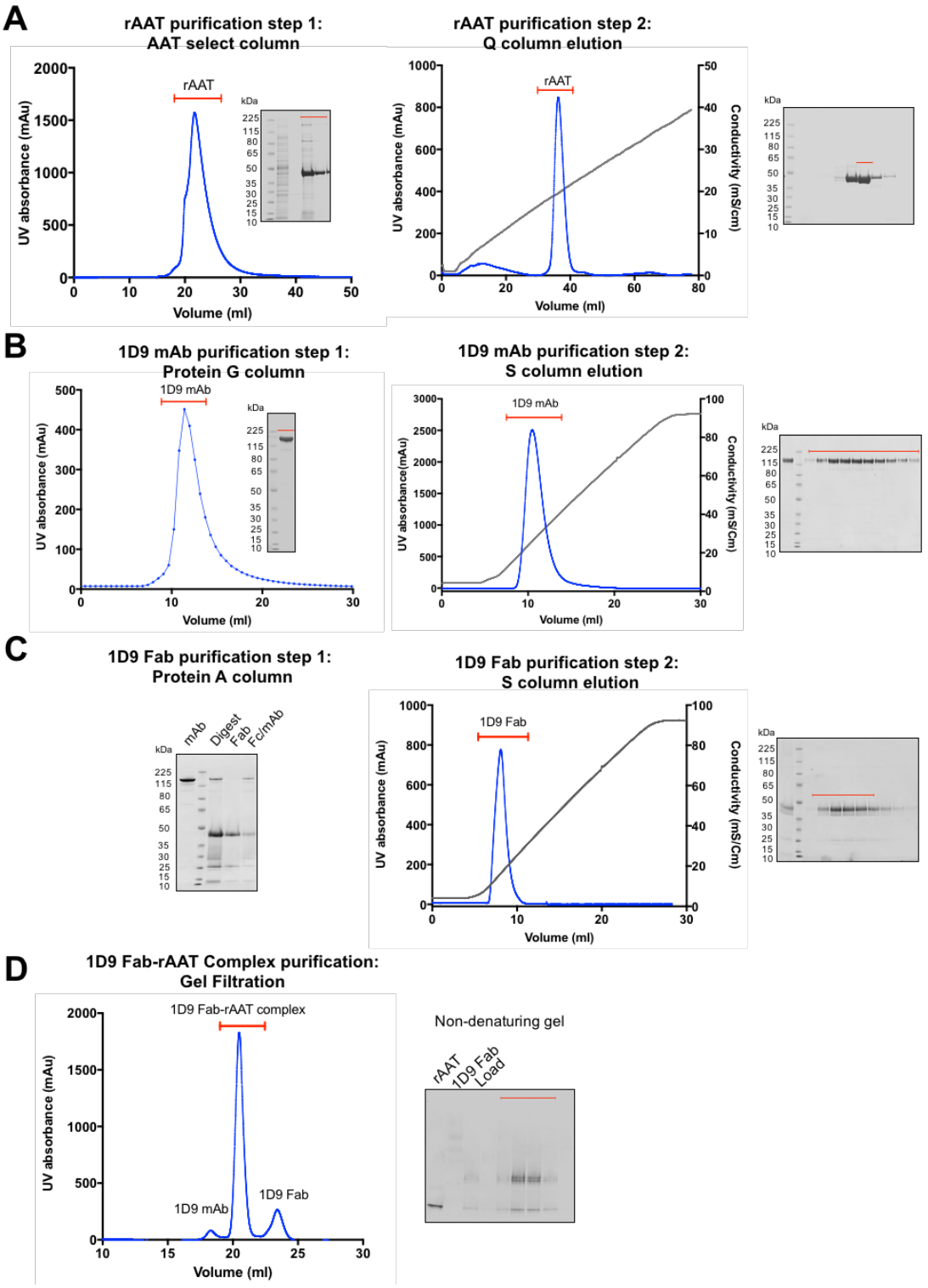


Figure 4.16. Preparation of samples for crystallographic studies. *UV absorbance traces of elution and denaturing gels of pooled elution fractions (red) are shown. A) Recombinant AAT was purified by elution from the Alpha$_1$-Antitrypsin Select column followed by elution from a Q-column. B) 1D9 mAb was purified by elution from a Protein G column followed by elution from a S-column. C) 1D9 Fab was generated by Ficin digestion and purified from the*

flow-through of a Protein A column. It was further purified by elution from an S-column. D) The rAAT-1D9 Fab complex was incubated for two hours and purified by size-exclusion chromatography. Fractions were run on a non-denaturing gel to assess complex formation (marked by the red line), next to controls of rAAT and 1D9 Fab. The sample prior to purification was also run (load).

The three samples - 1D9 Fab-rAAT complex, 1D9 mAb and 1D9 Fab - were concentrated to 10.2 mg/ml, 10.1 mg/ml and 5 mg/ml respectively in a low-salt buffer. They were subjected to sparse matrix crystallisation screens in a sitting-drop format by combining 100nl protein solution with 100nl precipitant solution and with vapour diffusion against 80µl of the latter buffer. The screens trialled were ProPlex (Molecular Dimensions), Structure Screen 1+2 (Molecular Dimensions), PACT (Molecular Dimensions), Index screen 1 (Hampton Research), Salt Rx A1 (Hampton Research) and JCSG Screen A1 (Qiagen). These screens cover a broad range of previously successful crystallisation conditions for other proteins with a variety of buffers, pH, precipitants and additives.

The commercial screens showed a good variety of drop appearances, ranging from clear to light or heavy precipitate, indicating that the concentrations of the proteins were suitable for crystallography. The 1D9 Fab-rAAT complex commercial screens contained one condition where crystals formed after 12 days and continued to grow for another 26 days (Table 4.1). The crystals were long and needle-like. The condition was 0.1M sodium cacodylate pH 6.0, 15% w/v PEG 4000. A larger-scale hanging drop screen was set up around this condition, with equilibration of a 0.5µl:0.5µl protein: precipitant solution against 500µl of the latter, varying the concentration of PEG 4000 (10% to 20% w/v in 1.25% intervals) and the pH of sodium cacodylate (pH 5 to pH 7 in 0.5 unit intervals) as in Figure 4.6. Crystals formed in multiple conditions, favouring PEG 4000 concentrations of 13.75%-16.25% w/v and pH values of 6-6.5 (Table 4.2). These crystals formed after around 7 days but were very fragile. Protein solubility varies with temperature and so equivalent crystal trays were also set up at 4℃. These crystals took longer to form but some appeared thicker.

The 1D9 mAb did not produce any crystalline hits in the commercial screens. This is not surprising: there are a lack of mAb crystal structures in the PDB, most likely due to the flexible loops and varied domain conformations of the Fab and Fc regions[212] and glycosylation of the Fc region.

The 1D9 Fab produced multiple crystalline hits in the screens and encouragingly, they were similar to conditions in different screens (Table 4.1). The conditions generally used PEG 3350 as the precipitant, with magnesium chloride and covering a range of pHs. The crystals either formed in clusters of long needles or were very intercalated, indicating multiple lattices could form from single nuclei. Crystal nucleation could be observed after one day. Custom hanging drop screens, like that shown in Figure 4.5, were again set up around the lead conditions, varying PEG 3350 concentration, magnesium chloride concentration and pH to produce bigger crystals. Intercalation and over-nucleation was still apparent in some wells but better-formed crystals were also apparent (Table 4.2). The concentration of 1D9 Fab subjected to crystallisation was also halved to attempt to reduce the rate of nucleation and intercalation. However, this appeared to not make a significant difference to the crystals that formed.

Screen	Well	Precipitant	Buffer
1D9 mAb			
/	/	/	/
1D9 Fab			
Structure	C9	30% w/v PEG 4000	0.2M $MgCl_2$, 0.1M Tris-HCl pH 8.5
JCSG	H11	25% w/v PEG 3500	0.2M $MgCl_2$, 0.1M Bis-Tris pH 5.5
Index	G10	25% w/v PEG 3500	0.2M $MgCl_2$, 0.1M Bis-Tris pH 5.5
Index	G11	25% w/v PEG 3500	0.2M $MgCl_2$, 0.1M Bis-Tris pH 6.5
Index	G12	25% w/v PEG 3500	0.2M $MgCl_2$, 0.1M Hepes pH 7.5
Index	H1	25% w/v PEG 3500	0.2M $MgCl_2$, 0.1M Tris-HCl pH 8.5
Index	H12	30% w/v PEG MME 2000	0.15M Potassium Bromide
1D9 Fab-rAAT complex			
ProPlex	B9	15% w/v PEG 4000	0.1M Sodium Cacodylate pH 6

Table 4.1. Conditions in the commercial screens tested that produced crystals for the three samples.

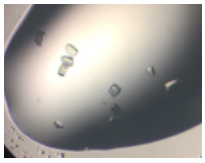
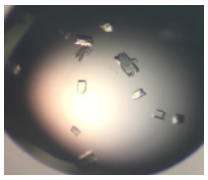
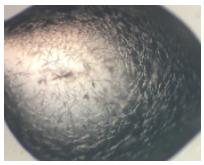
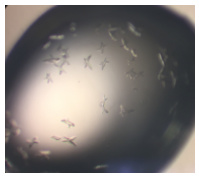
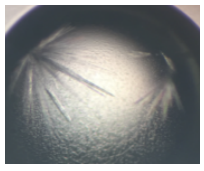
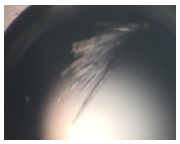
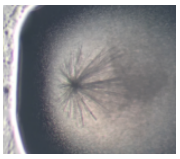
Precipitant	Buffer	Image
1D9 Fab		
20% PEG 3350	0.1M $MgCl_2$, 0.1M Bis Tris pH 6.5	
25% PEG 3350	0.2M $MgCl_2$, 0.1M Bis Tris pH 6.5	
30% PEG 3350	0.2M $MgCl_2$, 0.1M Hepes pH 7.5	
15% PEG 3350	0.2M $MgCl_2$, 0.1M Bis Tris pH 6.5	
1D9 Fab-rAAT complex		
15% PEG 4000	0.1M Sodium Cacodylate pH 6.0	
15% PEG 4000	0.1M Sodium Cacodylate pH 6.5	
16.25% PEG 4000	0.1M Sodium Cacodylate pH 6.5	
15% PEG 4000	0.1M Sodium Cacodylate pH 6 4°C	

Table 4.2. Examples of the crystals formed for both the 1D9 Fab and 1D9 Fab-rAAT complex in the hanging drop crystallisation screens.

Prior to data collection, some of the larger 1D9 Fab crystals were tested on the in-house X-ray source for diffraction. This was to confirm the crystals contained protein

and most usefully to test different cryo-protectants for the freezing of the crystals. Previous crystal structures of Fabs had used a variety of cryo-protectants ranging from 10-20% v/v ethylene glycol, 10-20% v/v glycerol and 15-25% w/v PEG 400 additions to the crystal reservoir solution[213-215]. When crystals of the 1D9 Fab were soaked using either 10% v/v ethylene glycol or 10% v/v glycerol as cryoprotectants, hairline fractures and cracks started to form in the crystals indicating instability. The crystal appearance did not visibly alter when soaked in the cryoprotectants 20% w/v PEG 400 and 20% w/v MPD conditions and so the majority of crystals were frozen in these conditions. Some crystals were also directly frozen into liquid nitrogen as the reservoir conditions had a relatively high level of PEG 4000 or PEG 3500 (≥15% w/v).

4.2.8 Determination of the 1D9 Fab structure

X-ray diffraction data were collected for the 1D9 Fab as detailed in methods 4.2.4.6. The highest resolution dataset came from a single crystal grown from equilibration against 0.05M $MgCl_2$, 20% w/v PEG 3350 and 0.1M Hepes pH 6.5. The crystal was 120µm by 80µm by 30µm in size and was frozen without any extra cryo-protectant (Figure 4.17A, B). This may explain why there are visible water rings in the diffraction pattern (Figure 4.17C).

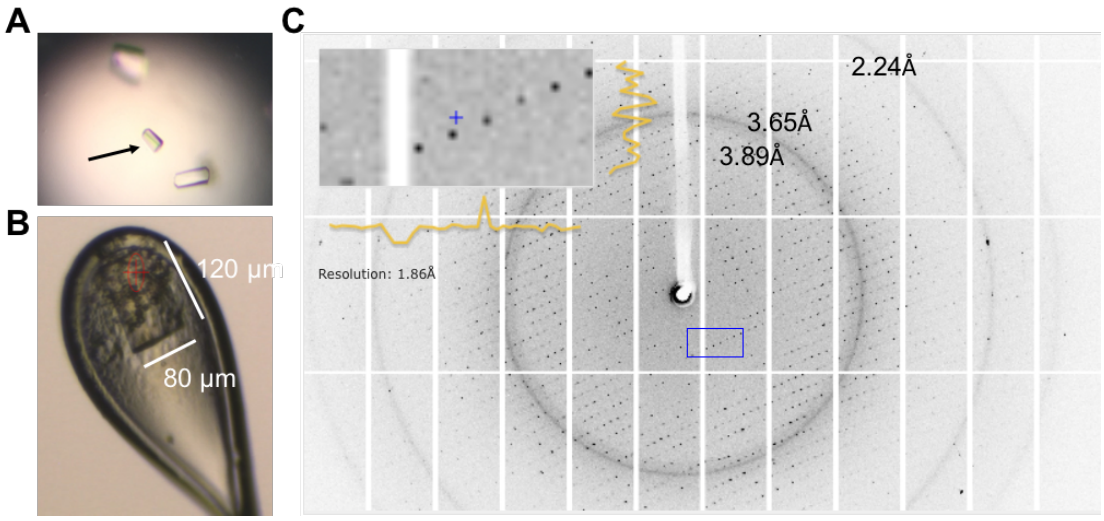


Figure 4.17. Data collection for the 1D9 Fab. *A) and B) show the crystal in the drop and on the cryoloop respectively. The crystal was 120µm by 80µm by 30µm. C) shows one diffraction pattern with spots apparent past the water rings at 3.89Å, 3.65Å and 2.24Å.*

Recorded data were processed and assessed for quality on the beamline using FastDp software[216]. Data were recorded in the resolution range 27.05-1.85Å with an overall 95% completeness. X-triage[217] highlighted one ice-ring related warning at 2.24Å. An initial space group of P 2 2 2 was obtained.

The merged data were used as input against a representative Fab structure (PDB ID: 1MF2[218]) for molecular replacement using CCP4 Phaser[219]. The 1MF2 Fab was also a IgG1 kappa chain Fab and had shown high sequence identity (97%) to the 1D9 light chain from initial sequencing results (Figure 4.18B)[218]. One molecule was found in the unit cell and the space group was revised to P 2 21 21.

To aid with structural determination, 1D9 was fully sequenced. The encoding RNA was isolated from the hybridoma cell line and converted to cDNA by reverse transcription using a polyT primer against the polyA transcription sequence. The individual chains were then amplified by PCR using forward degenerate primers and the poly T reverse primer. Specific amplification was confirmed by running the PCR products on an agarose gel before sending for DNA sequencing (Figure 4.18A). A lower sequence identity was found between 1MF2 and the 1D9 heavy chain (83%, compared to 97% for the light chain). Sequence modifications for both chains were made during the structure refinement process, whilst considering the observed electron density. At three residues, the 1MF2 sequence was modelled instead of the 1D9 sequence, based on a better fit with the electron density. These residues were Met→Leu (1D9→1MF2) at light chain position 4, Glu→Asp at heavy chain position 1 and Glu→Val at heavy chain position 5. The electron density was at a high enough resolution to allow their unequivocal identification with respect to known sequence propensity (Figure 4.18D). The sequence mismatches may be the product of sequencing errors as they are at the start of the chains and near the binding site of the degenerate primers. The sequence of the 1D9 heavy chain was also incomplete towards the conserved C-terminus of the chain; these residues were also unambiguously identified from comparison with known antibody sequences based on the electron density.

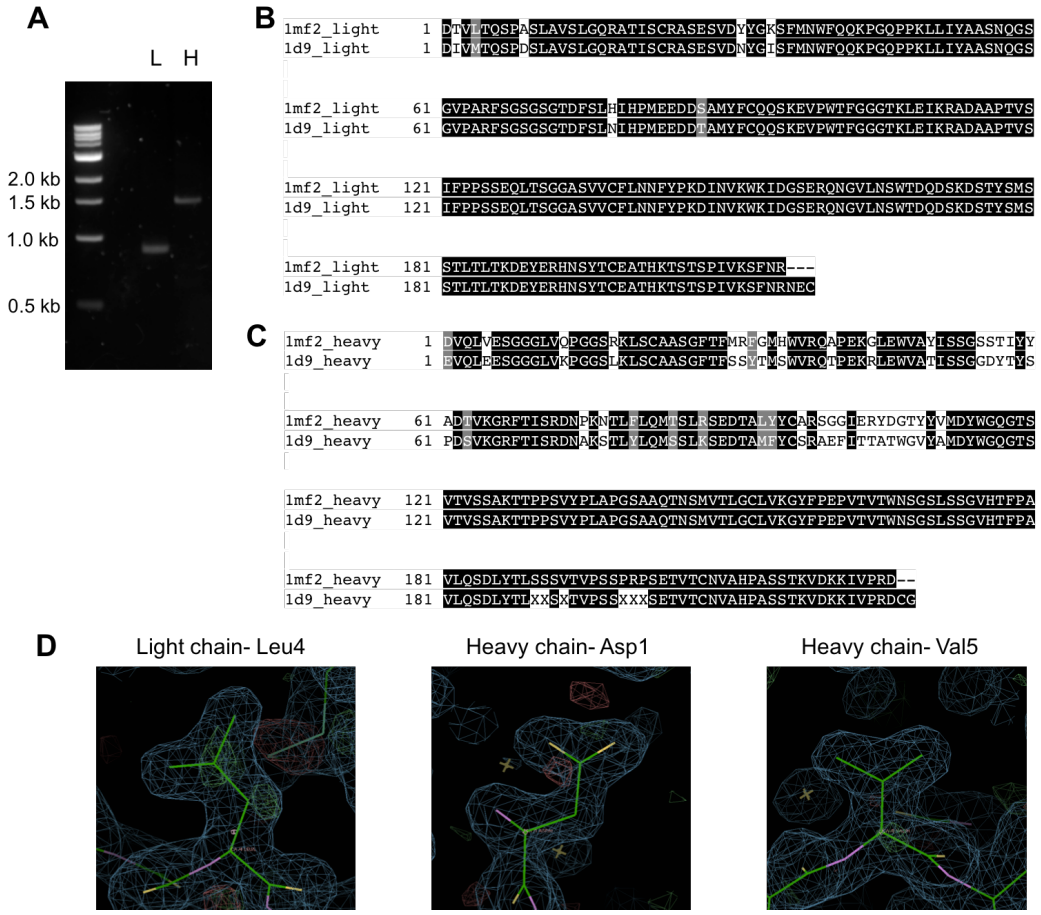


Figure 4.18. Sequencing of 1D9. *A) A 2% w/v agarose gel showing PCR amplification of the light chain (L) and heavy chain (H) of 1D9 using degenerate forward primers and a polyT primer. B) and C) A sequence alignment between 1MF2[218] (a IgG1 kappa Fab) and 1D9. 1MF2 was used as a model for molecular replacement. D) Electron density of residues that retained the 1MF2 sequence over the sequence identity proposed from sequencing of 1D9.*

The crystal structure of the 1D9 Fab was solved to 1.85Å resolution and is shown in Figure 4.20. Restrained refinement was performed using CCP4 Refmac5[220] and Phenix[221]. Structural visualisation and manipulation was performed using Coot[161]. During the refinement process, there were spikes in both the R value and R free in resolution shells that coincided with ice rings, reflecting poor data quality in these regions. These narrow slices of data were therefore excluded from consideration.

In the final refined structure, there were 3880 atoms comprising two protein chains (the heavy and light chain), 584 unique water molecules (589 including alternate conformations) and 1 magnesium ion (Table 4.3). The light chain consisted of 216 residues and could be fully built into the density whereas there was missing electron

density for 6 residues (139-144) out of the 225 residues comprising the heavy chain. Some residues were not resolved at both chain C-termini. There was also missing density for 23 surface-exposed side chains, which are spread throughout the structure. Most of these residues are charged or have long side chains (Lys, Arg, Glu, Asp, Gln, Asn) suggesting that are most likely not resolved due to flexibility or radiation damage[222]. A Ramachandran plot shows 99% of residues are in energetically favoured regions, with the remaining 1% in energetically allowed regions (Figure 4.19). There are no rotamer outliers, 1.8% RSRZ outliers (real-space R-value Z score, a measure of the fit to the density) and a low clash score of 2. The overall structure has a B-factor of 19, which corresponds to an uncertainty of around 0.5Å per atom[124]. This is considered to reflect low disorder within the crystal and is visually represented by a high proportion of blue in the coloured heat map in Figure 4.21.

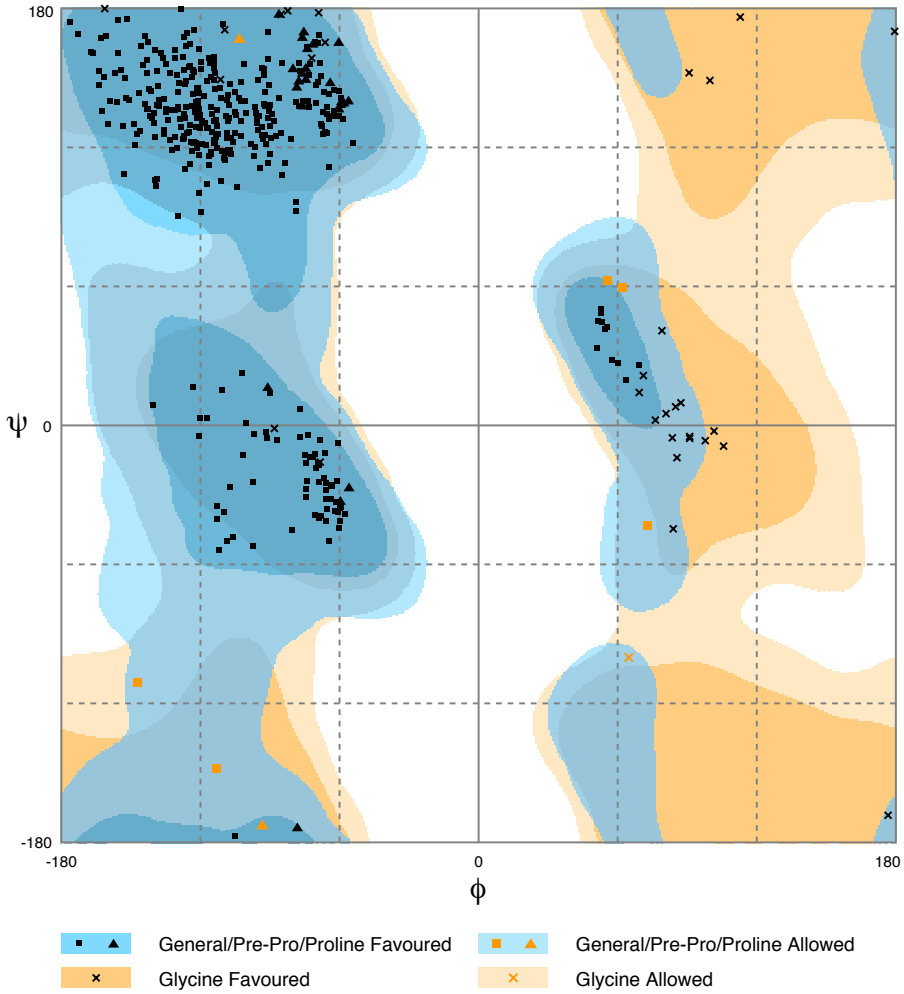


Figure 4.19. A Ramachandran plot of the 1D9 Fab structure. *Nearly 99% of residues are in energetically favoured regions and the remaining 1% in energetically allowed regions. Plot made using RAMPAGE*[223].

Wavelength ()	1.0
Resolution range (Å)	27.05 - 1.852 (1.919 - 1.852)
Space group	P 2 21 21
Unit cell a, b, c (Å)	54.103 85.257 89.427
α, β, γ (°)	90 90 90
Total reflections	431326 (35044)
Unique reflections	34101 (2951)
Multiplicity	12.6 (11.9)
Completeness (%)	94.71 (83.82)
Mean I/sigma(I)	16.16 (2.90)
Wilson B-factor ($Å^2$)	19.06
R-merge	0.1191 (0.8111)
Reflections used in refinement	34003 (2938)
Reflections used for R-free	1718 (155)
R-work	0.1860 (0.2986)
R-free	0.2221 (0.3724)
Number of non-hydrogen atoms	3880
macromolecules	3290
ligands	1
solvent	589
Protein residues	435
RMS(bonds)	0.005
RMS(angles)	1.04
Ramachandran favored (%)	98.83
Ramachandran allowed (%)	1.17
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clash score	2.19
Average B-factor ($Å^2$)	19.96
macromolecules	18.29
ligands	31.57
solvent	29.29

Table 4.3. Data collection and refinement statistics for 1D9 Fab. *Table generated from the Protein Data Bank (PDB) validation server, deposited PDB ID: 6GXX. Statistics for the highest-resolution shell are shown in parentheses.*

The 1D9 Fab is composed of two variable domains (V_L and V_H) and two constant domains (C_L and C_H1), which each follow the typical β-sandwich immunoglobulin fold. The fold is stabilised by four pairs of disulphide bridges (in the heavy chain: Cys22-Cys96 and Cys152-Cys207, in the light chain: Cys23-Cys92 and Cys138-Cys198) according to the Kabat numbering scheme[224]. All but Cys23-Cys92 showed alternate conformations within the lattice owing to discrete disorder. The cysteines that are involved in the inter-chain disulphide between the heavy chain and light chain in the Fab region (Cys227 and Cys218 on the heavy and light chain respectively) were not resolved in the structure.

The proposed magnesium ion is hydrogen-bonded between Tyr144 and Tyr177 on the light chain and several water molecules in octahedral geometry. The magnesium is unlikely to have a functional or structural role in the Fab but merely present because of the composition of the crystallisation buffer.

The complementarity-determining regions (CDRs) form the antigen binding site and in the 1D9 structure are well-determined. They are labelled L1 to H3 and were identified based on a consensus between the Kabat and Chothia numbering schemes[224]. H3 is generally thought to be the most important CDR for antigen binding and hence is the CDR that has the biggest variation in length between antibodies (between 3-25 residues)[225,226]. In the case of 1D9, H3 is 16 residues from Ala99 to Thr114. The CDRs have larger B-factors than the rest of the structure (up to a maximum of 58.5), particularly H3 (Figure 4.21). This indicates a high atomic displacement of atoms, reflective of atomic mobility and flexibility of the chain. The CDRs are not involved in extensive crystal contacts probably contributing to their disorder. Half of the RSRZ outliers, where the RSRZ is a measure of the geometric fit between a residue and the electron density, are in H3 from Thr103-Thr106. The poorly resolved electron density in this region again can reflect flexibility as small differences between copies of the protein in the crystal lattice will tend to ‘blur’ the density. The other RSRZ outliers are residues at the chain break in the heavy chain or the C-termini. In these regions, there are similarly high B-factors which helps to explain why electron density was not apparent for the missing residues. Lastly, electrostatic potential surface analysis shows the antigen binding surface is weakly negatively charged which may have a role in electrostatic interactions between 1D9 and AAT.

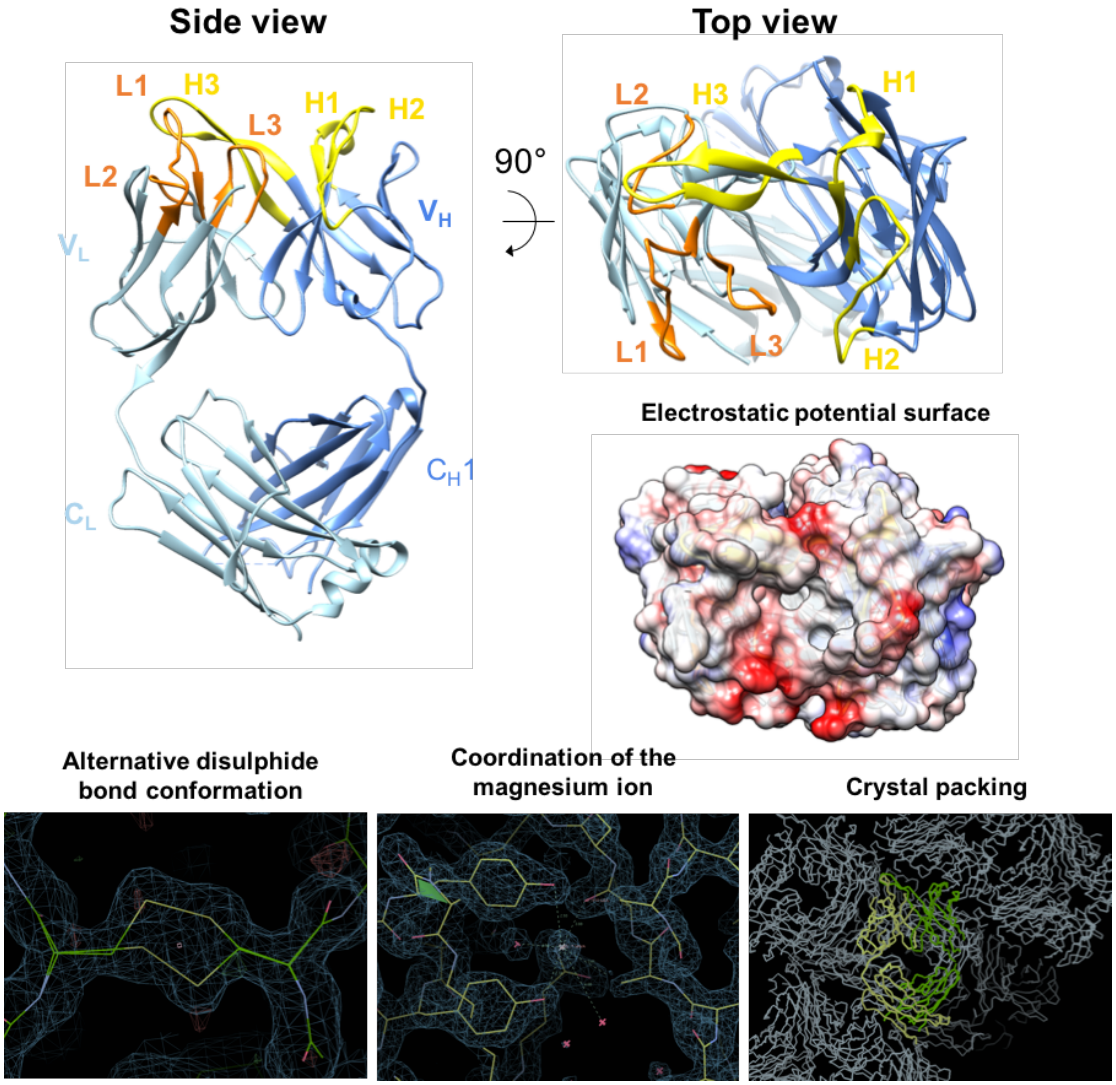


Figure 4.20. The 1D9 Fab structure. *A side and top view of a cartoon representation of the 1D9 Fab made using Chimera[9]. The heavy chain (H) is coloured in dark blue and the light chain (L) in light blue. The constant and variable domains are as labelled. The 3 heavy chain CDRs are coloured in yellow and the 3 light chain CDRS are coloured in orange. The electrostatic potential surface shows the negative charge for the antigen-facing side of the Fab. The panel at the bottom shows an example of an alternative disulphide bond conformation (left), coordination of the magnesium ion (centre) and the crystal packing of the 1D9 Fab (right), visualised using Coot[161].*

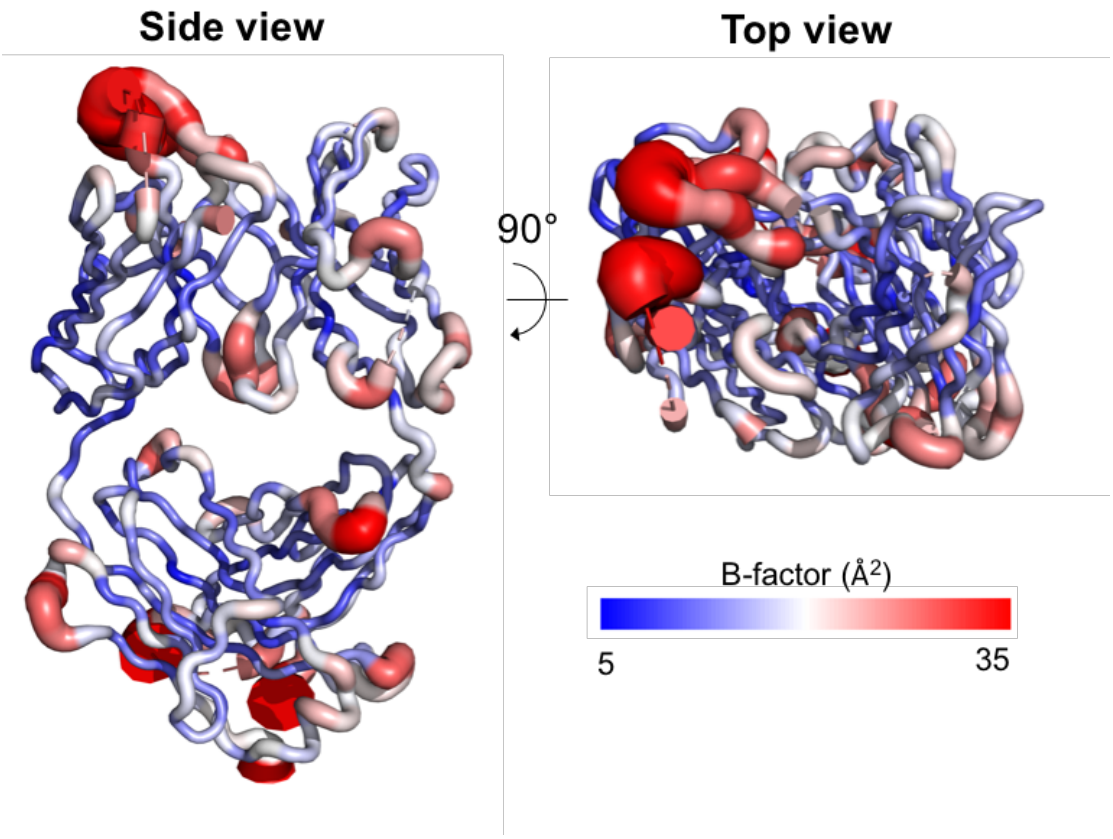


Figure 4.21. B-factor analysis of the 1D9 Fab. *The Fab has been coloured from blue to white to red together with increased chain thickness in order of increasing B-factor values using PyMOL*[163]*. A side and top view are shown and indicate that areas of high B-factors are in the loops, particularly the CDRs and at chain breaks/termini.*

4.2.9 The structure of 1D9 Fab in complex with rAAT

Diffraction data were collected as detailed in methods 4.2.4.6, on around 20 crystals, which typically showed diffraction to around 5-11Å, making them unusable for structure determination. A single plate-like crystal of the 1D9 Fab-rAAT complex measuring 280µm by 30µm by 20µm in size, yielded diffraction to better than 3Å. The crystal was grown from equilibration against 15% w/v PEG 4000, 0.1M sodium cacodylate pH 6 and was frozen in 20% w/v MPD as a cryoprotectant. In comparison to 1D9 Fab alone, no ice rings were observed in the diffraction pattern (Figure 4.22).

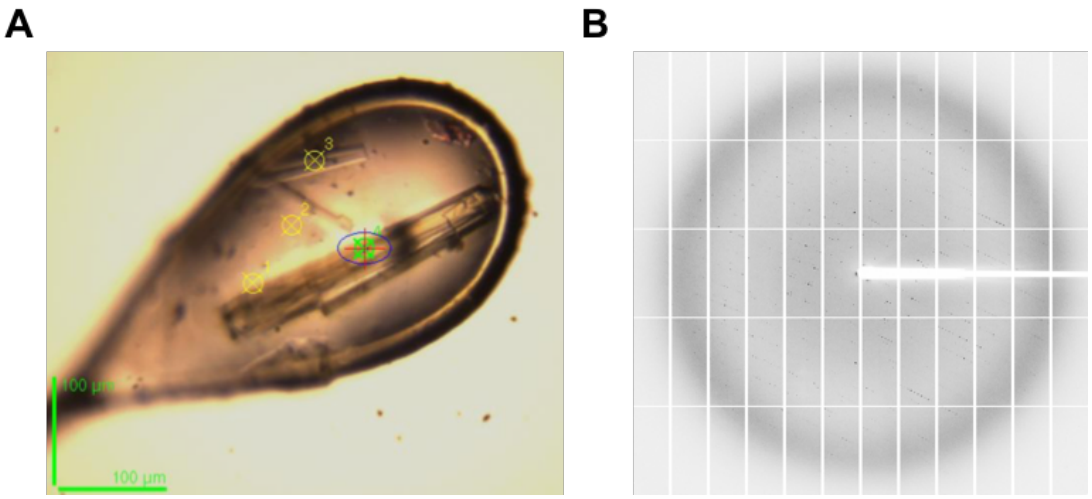


Figure 4.22. Data collection for the AAT/1D9 Fab complex. *A) The crystal of the complex on the cryoloop with the collection area highlighted in green. B) One diffraction pattern of the crystal.*

Recorded data were processed using XDS to yield an initial space group of C 1 2 1. Data were recorded in the resolution range 49.2-2.9Å with an overall 99% completeness. X-triage[217] did not detect any ice-ring related problems. The data appeared moderately anisotropic. Molecular replacement was run using CCP4 Phaser[219] with native AAT (PDB ID:1QLP[8]) and the previously solved 1D9 Fab (PDB ID:6GXX) as search models[8]. Two molecules of each (i.e. two complexes) were present in the asymmetric unit.

The final structure was refined to 2.95Å (Figure 4.23). Restrained refinement was performed using CCP4 Refmac5[220] and Phenix[221] and the structure was built using Coot[161]. Due to the low resolution, non-crystallographic symmetry restraints were used in the refinement process and refinement of two B-factors per residue was attempted but discarded in favour of one B-factor per atom. Translation/Libration/Screw (TLS) rigid-body descriptions were also tried during refinement, but as they did not seem to make a difference to the quality of the structure, the final model was refined in their absence. The statistics for the integration of the diffraction data were representative of a high-quality dataset, yet one copy of the complex showed fairly extensive gaps in the density. Crystal defects such as twinning, pseudosymmetry and mis-indexing were excluded as reasons for this. The solvent content of the crystal was close to the lower end of the range typically observed for protein crystals, and it may be that the high packing density

resulted in molecules constrained in slightly different configurations resulting in blurring of the density.

Wavelength (Å)	1.0
Resolution range (Å)	49.2 - 2.95 (2.95- 3.00)
Space group	C 1 2 1
Unit cell a, b, c (Å)	247.48 36.75 201.78
α, β, γ (°)	90 126 90
Total reflections	102568 (4738)
Unique reflections	31780 (1472)
Multiplicity	3.2 (3.2)
Completeness (%)	99.1 (93.3)
Mean I/sigma(I)	10.3 (1.4)
Wilson B-factor (Å^2)	66.8
R-merge	0.095 (0.894)
Reflections used in refinement	30,277
Reflections used for R-free	1503
R-work	0.233
R-free	0.282
Number of non-hydrogen atoms	11031
Protein residues	1,503
RMS(bonds)	0.25
RMS(angles)	0.48
Ramachandran favored (%)	95
Ramachandran allowed (%)	5
Ramachandran outliers (%)	0
Rotamer outliers (%)	2
Clash score	6
Average B-factor (Å^2)	77.0

Table 4.4. Data collection and refinement statistics for the AAT/1D9 Fab complex. *Table generated from the preliminary Protein Data Bank (PDB) validation server. Statistics for the highest-resolution shell are shown in parentheses.*

There was a total of 11,031 atoms comprising four unique molecules and six protein chains (two AAT chains, two 1D9 heavy chains and two 1D9 light chains) (Table 4.4). The AAT chains consisted of 350 and 332 residues each, with three and six chain breaks respectively. Particularly for the latter AAT molecule, this is a high level of missing information, as mentioned above. These chain breaks and missing residues were throughout the structure (residues 103-108, 120-130, 136-140, 173-179, 312-325 and 346-353). The heavy 1D9 chains consisted of 222 and 210 residues – each with one chain break – and the light 1D9 chains were complete with 217 and 212 residues. Missing residues in the Fabs were mainly at the chain termini.

Ramachandran analysis showed 95% of residues are in energetically favoured regions and 5% in allowed regions. The structure has a rotamer outlier score of 2% and clash score of 6.

The orientation of the molecules clearly identified the epitope that is recognised by 1D9 Fab. 1D9 Fab binds to the top of the AAT molecule, interacting with the C-terminal end of the RCL, β-sheet C (mostly s1C and s2C) and part of β-sheet B (s1B and s2B) of AAT. This was in agreement with the results from the earlier epitope mapping experiments (4.3.5). The interaction between 1D9 Fab and AAT formed an average solvent-accessible surface of 888Å^2 between the two copies of the complex. This was calculated in Chimera by generating the molecular surfaces for the complex and for the individual components, and dividing the difference by two[9]. This is in the typical range of interface region surface areas (1103 ± 244Å^2) between antibodies and their antigens[227].

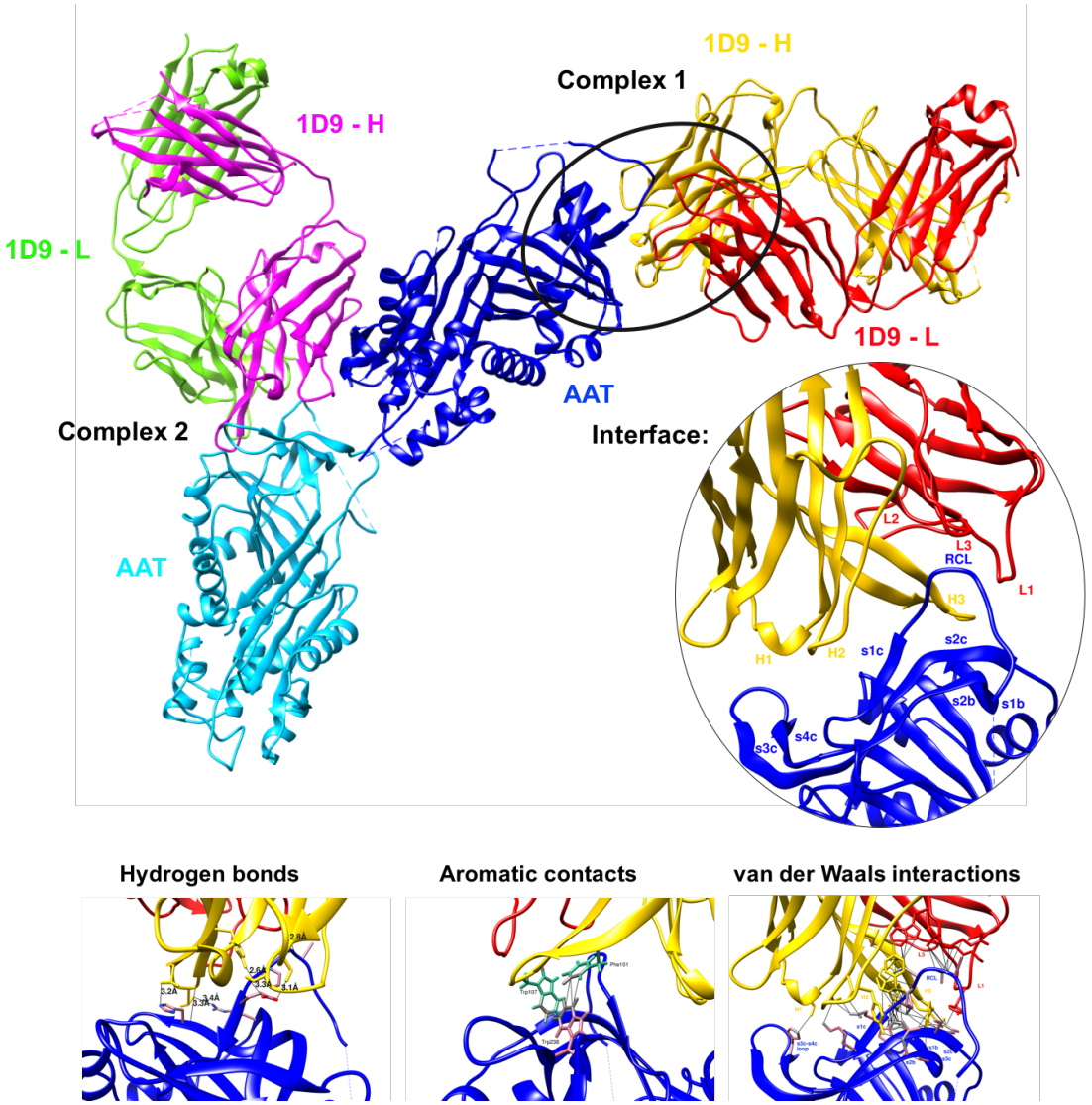


Figure 4.23. The crystal structure of the 1D9 Fab-rAAT complex. *Two copies of the complex were present in the asymmetric unit (top) as displayed in cartoon format using Chimera[9]. AAT molecules are coloured in blue and cyan, the 1D9 Fab heavy chains (1D9-H) in yellow and magenta and the 1D9 Fab light chains (1D9-L) in red and green. The insert shows a zoomed-in interface between AAT and 1D9 in complex 1, highlighting the CDRs and the areas of interaction in AAT. The panel at the bottom shows the interactions between AAT and 1D9 in complex 1. Hydrogen bonds, aromatic contacts and van der Waals interactions were all calculated using Chimera[9].*

The structure analysis tool in Chimera was used to look at the specific interactions between AAT and 1D9 Fab[9]. This was carried out for the complex with better-resolved RCL residues to ensure that the interaction analysis was more complete (complex 1). All of 1D9's CDRs apart from L2 are involved in antigen recognition of AAT, however most contacts are made by residues in H2, H3 and L3. Since the

model was built without hydrogen atoms, hydrogen bonds were inferred when a donor and acceptor were within 3.5Å of each other[228]. A total of seven hydrogen bonds were found between the two molecules with distances between 2.6Å and 3.4Å (Figure 4.23). Four of these involved H3 residues hydrogen bonding to residues in the RCL, s1B and s2C (Thr103-His231, Thr106-His231, Thr106-Arg282 and Tyr110-Pro361). The other hydrogen bonds were made from Ser52 and Ser53 in H2 both to Glu363 in s1C and Val98 in L3 to Met358 in the RCL. One aromatic contact was identified involving Phe101 and Trp107 in H3 and Trp238 in s2B. No salt bridges were identified and electrostatic potential surface analysis of the top surface of AAT showed there were areas of both positive and negative charge. A total of 76 van der Waals contacts between atoms were recognised. In total the number of residues involved in any interaction were as follows: for 1D9, H1:1, H2:5, H3:5, L1:1, L2:0, L3:4, and for AAT, RCL:5, β-sheet C:6, β-sheet B:3. Of the residues in the 1D9 paratope, Tyr and the other aromatic residues dominated the interactions, which has been observed in a systematic review of antigen-antibody interaces[227]. 1D9 Fab did not interact with AAT using residues from any of the framework regions that flank the CDRs in keeping with their primary role as a scaffold[226].

PyMOL was used to assess the structural differences between molecules by calculating the root mean square deviation (RMSD). The two complexes in their entirety within the asymmetric unit were first compared (Figure 4.24). The superimposition of the two complexes yielded a RMSD of 0.8Å. The AAT molecules superimposed well with RMSD of 0.5Å, indicating high structural similarity. The major areas of variation between the two AAT structures are a shift of helix D and variations in the chain position in loops or at chain breaks. More specifically the latter included the loop between helix F and s1A, the loop between helix F and s3A, the chain break before s4A and the chain break in the RCL. The variable domains of the Fabs also superimposed well, including all loops, but there was more variation in the constant domains. This indicates there is relative flexibility between the constant and variable domains of the Fab as has been seen before in other Fab structures[168].

Despite their structural similarity, there was a large difference in the B-factors between the complex structures as highlighted by the heat map in Figure 4.25. In complex 1, all Fab residues had low B-factors, whereas in complex 2, the Fab constant domain generally had high B-factors. Of note however, is that the CDRs in both complexes had low B-factors, which contrasts with the high B-factor of H3 in the

unbound Fab structure (Figure 4.21). For AAT, helix D and F had the highest B-factors in complex 1 but these regions were stable in complex 2. Instead, in complex 2, β-sheet C was more mobile.

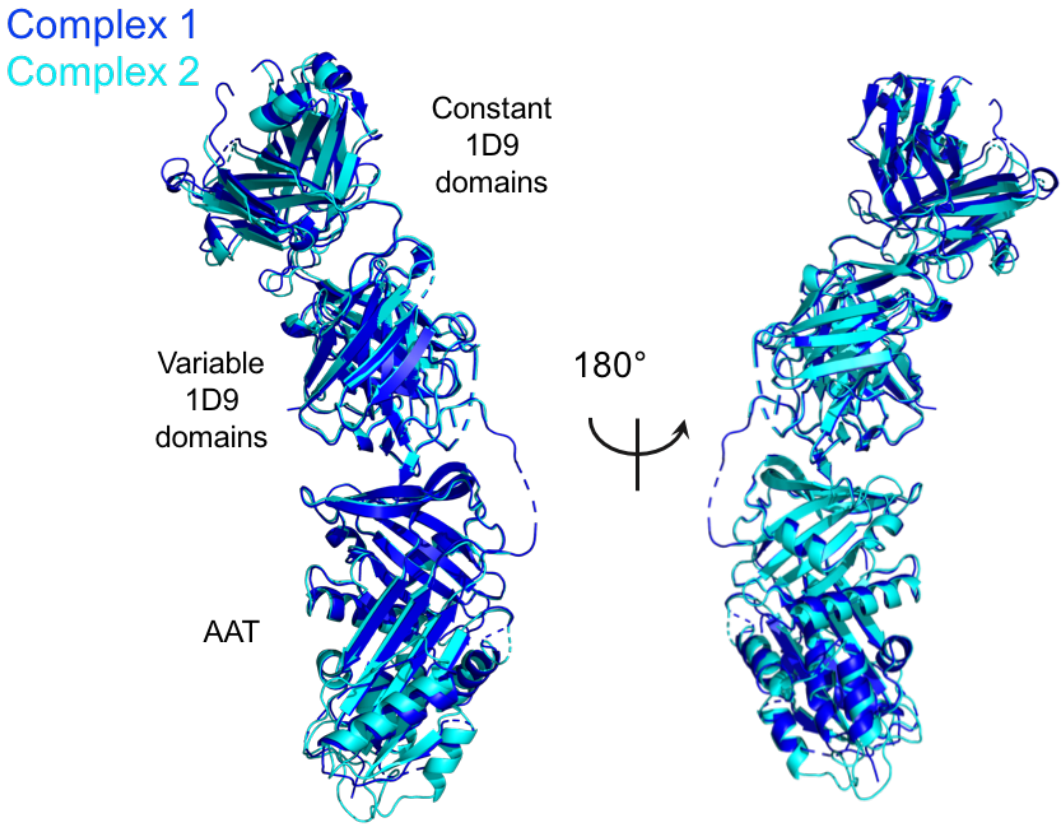


Figure 4.24. Overlay of the two complex structures present in the unit cell. *Complex 1 is shown in dark blue and Complex 2 in cyan. PyMOL was used to align the structures and calculate the RMSD*[163]*.*

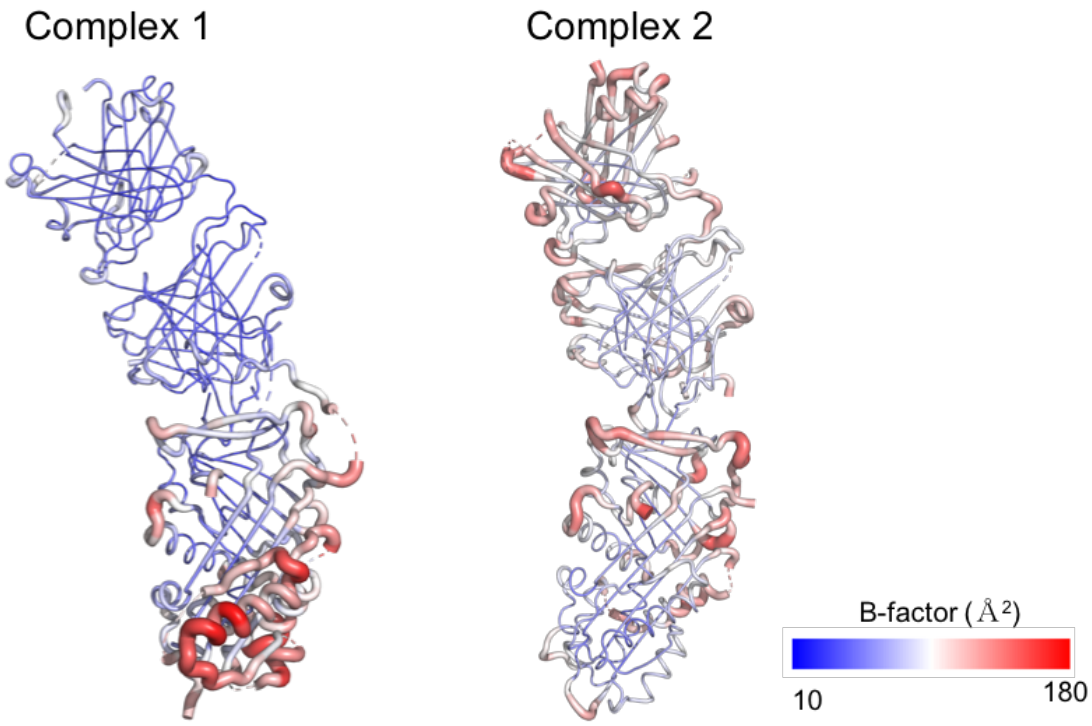


Figure 4.25. B-factor analysis of the two complexes present within the unit cell. *The molecules were coloured from blue to white to red, together with increased chain thickness, in order of increasing B-factor values in PyMOL*[163].

Complex formation causes structural variations in both antigen and antibody[229]. Therefore, the AAT structures in the complexes were also compared to previously solved crystal structures of native AAT. As the AAT structures within both complexes of the unit cell shared a relatively low RMSD and therefore were structurally similar, only the AAT in complex 1 was assessed. AAT was compared to the crystal structures of intact, native AAT (PDB ID: 1QLP and 3NE4)[8,230]. Both these previous native structures have an RMSD of 0.3Å between each other and hence 1QLP was used as a representative. The RMSD between 1QLP and the AAT of complex 1 was 0.9Å. There were 3 main areas of variation: a shift in helix D, a tilt of helix F and an altered conformation of the RCL (Figure 4.26). These areas also had the highest B-factor in complex 1, reflective of intrinsic movement and are known to be particularly mobile components of the serpin structure[1,231-233]. Their altered conformations may be a result of 1D9 binding to AAT[229] or simply conformational mobility influenced by

crystal packing. Allosteric changes have been noted for other serpins that transmit binding events between helix D and the RCL, in particular the breach region[234].

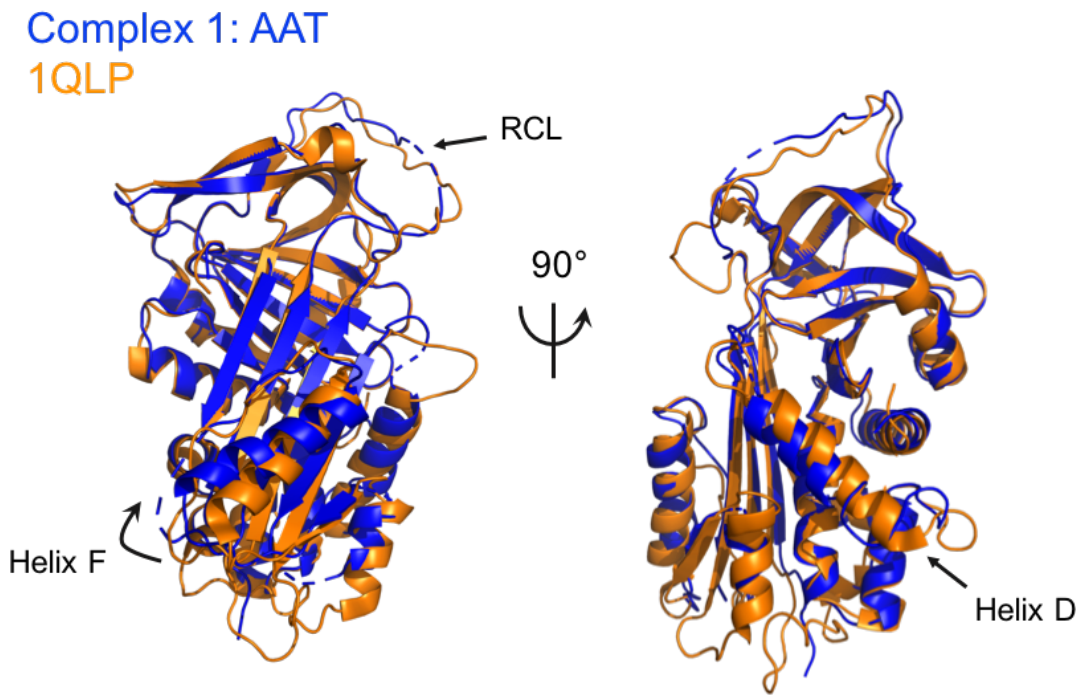


Figure 4.26. An overlay of AAT crystal structures. *In orange is 1QLP[8] and in dark blue is AAT from complex 1. The 3 areas (Helix D, Helix F and the RCL) of biggest variation are highlighted. Structural comparison and the RMSD calculation was carried out using PyMOL[163].*

Lastly, the 1D9 Fab structures in the unbound and antigen-bound states were compared to ascertain structural changes resulting from complex formation. The superimposition of the 1D9 Fab structure on the 1D9 Fab in the first complex showed a RMSD of 1.2Å. This indicates a small conformational change between the two states consistent with what is frequently observed for antibody interactions[229,235]. Deviations were observed at the chain termini, the flexible loops linking β-strands and in the CDRs. H1, H2 and particularly H3 all showed high RMSD relative to the rest of the structure. H3 has been documented to show significant binding-related conformational changes in a third of antibodies[168]. In the light chain, only L1 showed a significant chain deviation. The deviations and high superimposition of L2 and L3 between the structures can be seen in Figure 4.27. There were no obvious conformational changes in the relative orientations of the heavy and light chains, the elbow angle between the constant and variable domains nor the first loop in the CH1

domain, which has been observed in multiple antibody structural comparisons[168]. The superimposition of the Fab domains associated with each complex showed a RMSD of 1.4Å. Similar conformational shifts were observed for loops and again were highest for the CDRs, in particular H2, H3 and L1. The slightly higher RMSD between the complexes with respect to that seen with AAT is most likely due to the greater amount of missing density in complex 2; this leads to a greater uncertainty in model building at these areas.

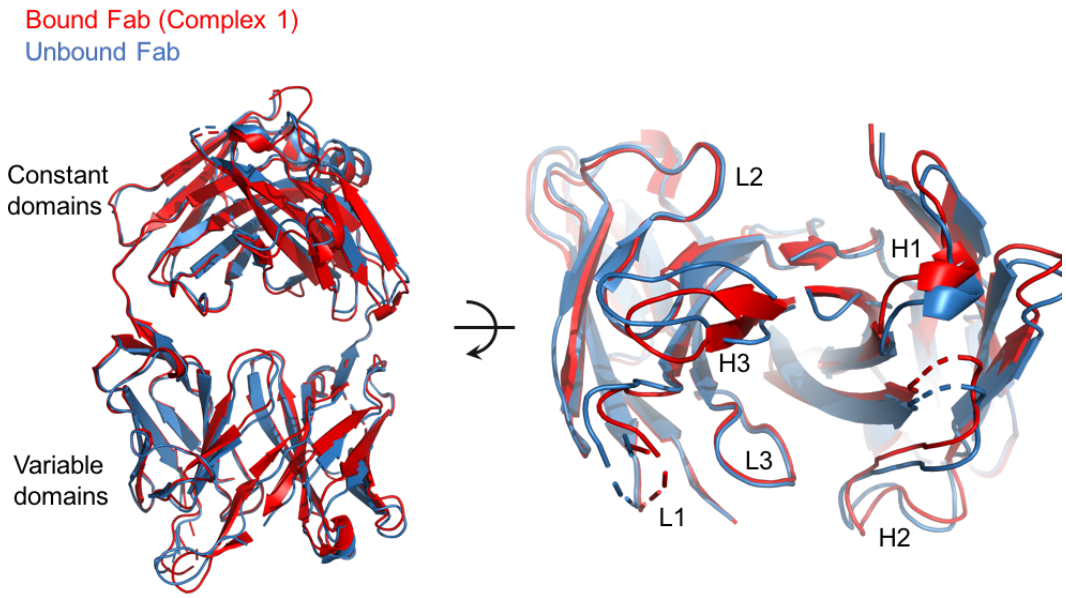


Figure 4.27. Comparison of 1D9 structure in unbound and bound states. *The unbound Fab structure is shown in blue and the bound Fab structure from Complex 1 is shown in red. Side and top views are displayed with the CDRs labelled. Structural comparison and RMSD calculation was carried out using PyMOL*[163].

4.2.10 NS-EM studies of 1D9 Fab bound to polymers

Although 1D9 has a higher affinity for the AAT monomer, it does still bind to AAT polymers. This is despite the 1D9 epitope being in a region of AAT that is intrinsically linked to polymerisation. This may provide clues as to the structure of the polymer. To characterise this interaction, purified polymer from the liver of a ZZ-AAT homozygote was incubated with 1D9 Fab and visualised by NS-EM (Figure 4.28). The appearance of the Fab-incubated polymer was not visually distinct from the polymer in its absence. This indicates that the 1D9 Fab is likely not binding to every

AAT subunit, as observed for 4B12 Fab (Figure 6.7). This is also in accordance with the 10-fold lower affinity for polymer than monomer calculated from the ELISA against different conformers. However, in the incubated polymer chains there are protrusions that are reminiscent of Fab-bound subunits (marked by yellow asterisks in Figure 4.28). This suggests a proportion of Fab can bind but that the epitope may be obscured in the polymer. The polymer sample was also incubated with a five-fold excess of Fab to see whether Fab labelling increased, however the appearance of polymer chains did not change (not shown). Unfortunately, the data were too inconclusive to compare to the equivalent experiment with heat-induced polymers to try and visually explain how heat polymers have a much lower recognition than liver polymers by 1D9.

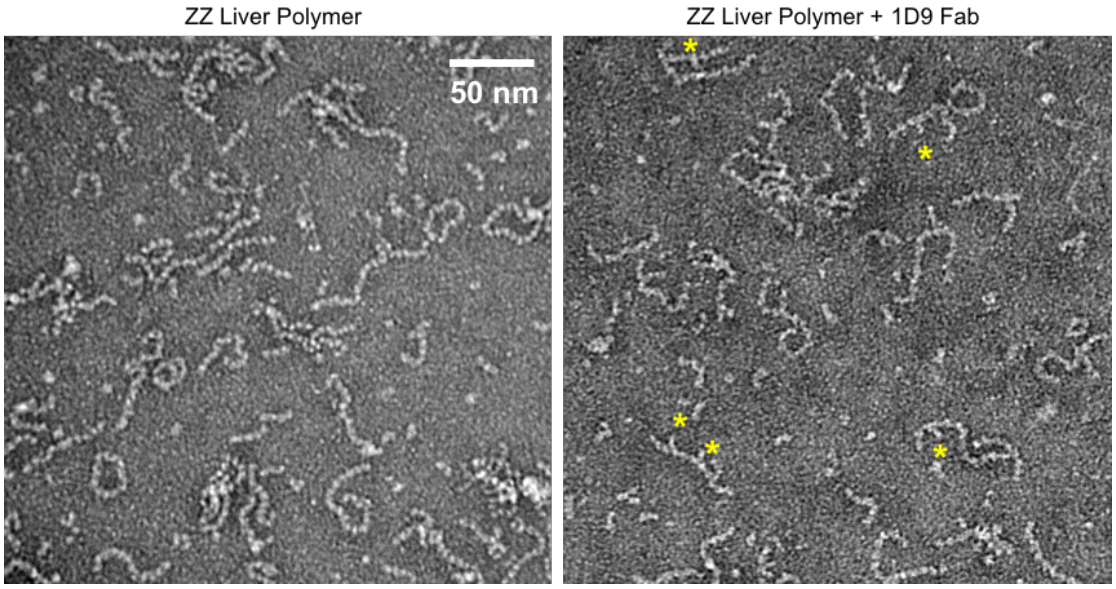


Figure 4.28. Negative stain electron micrographs of ZZ liver polymers in the absence and presence of 1D9 Fab. *The micrographs were imaged on a Tecnai 100kV with a CCD camera at nominal magnification of 64,000X. Scale bar of 50nm is shown. The yellow asterisks identify the Fab-bound subunits.*

4.4 Discussion

4.4.1 Monoclonal antibodies highlight the conformational differences between AAT species

The mechanism of AAT polymerisation and the structure of AAT polymers from different origins have not yet been fully determined. Conformation--specific antibodies represent tools that have been used to investigate these characteristics of AAT pathology. For example, the development of 2C1 demonstrated that heat-induced polymers and liver polymers share an epitope that is absent from urea- and acid-induced polymers and epitope mapping of the polymerisation-blocking antibody 4B12 revealed a propagation of conformational changes that occurs throughout the AAT molecule during polymerisation[76,116,117].

The usefulness of these antibodies in characterising the AAT polymerisation pathway led to the current project: the development of new AAT-specific mAbs that displayed novel or improved characteristics against different AAT antigens. Previously, mice were immunised with monomeric Z-AAT or heat-induced Z-AAT polymers[76,116]. However, in this hybridoma screen, with a view to produce antibodies with different reactivities, the immunising antigens were either Z-AAT polymers purified from patient plasma or a mix of acid- and urea-induced polymers. A total of 31 antibodies were developed and their specificity was characterised by a screening ELISA. They were categorised into five groups that showed distinctive selectivity against four polymeric species and two monomeric species. The breadth of the observed reactivities (preference for polymers over monomer, monomer over certain polymers, Z monomer over M monomer and M-based AAT species over Z-based AAT species) highlights the multiple structural differences that exist between different AAT conformers. Indeed, it is well documented that the monomer undergoes structural changes during polymerisation, that the Z mutation causes increased structural lability compared to M-AAT and that polymers vary structurally depending on their method of induction or origin[48-52,69-72,76]. These differences are also reflected by the running profile of AAT conformers on a non-denaturing gel[69,70].

Antibodies with similar reactivity to previously developed mAbs were generated in this screen in groups 1 (2C1-like) and 3 (9C5/4B12/3C11-like). The group 3 antibodies were not characterised further as they did not display novel reactivities. Secretion levels from a hybridoma cell line in group 1 (1A3) was estimated, to

determine whether a 2C1-like antibody could be identified that had a higher expression level but 1A3 levels were not significantly improved. Antibody production by hybridomas is an unstable phenotype and there is a natural variation in secretion levels of clones[179]. Culture conditions, particularly the growth media, have been shown to have an effect[236,237]. Based on hybridoma culture in the lab, antibody levels in the CMS can vary significantly from 1mg/100ml to 20mg/100ml. Loss of antibody productivity exhibited by hybridomas over time has been associated with growth of a non-producing subpopulation that no longer synthesises the heavy chain[236]. Low-producing lines can try to be recloned by limiting dilution to find high-producing clones[179], but this has not worked with 2C1.

Most interesting were the groups which displayed novel reactivity, such as group 4 (preference for monomer and plasma ZZ-AAT polymers) and group 5 (preference for M-AAT over Z-AAT). As with previous screens, there was not a direct correlation between the conformational state of the immunising antigens and the reactivity observed in the newly-generated antibodies. For example, no antibodies were generated that showed specific selectivity against the urea- or acid-induced polymers, which are believed to be structurally different[69], despite their use as antigens. Furthermore, 1D9 in group 4 displayed a strong preference for monomers despite mice being only immunised with polymers and yet no such antibody was generated when mice were immunised with only monomers (E. Miranda, personal communication). M-AAT specific antibodies have not previously been generated and they complement the ATZ11 antibody, which displays specificity against Z-AAT. This is the first published AAT antibody screen that used M-AAT-based antigens for immunisation, and group 5 antibodies showed preference for this wild-type variant.

4.4.2 Insights about AAT from the characteristics of 1D9

The 1D9 mAb was selected for in-depth characterisation based on its unique reactivity profile against AAT antigens. Insight into its epitope would provide some degree of structural information on the region that distinguishes the conformations it recognises from those it does not. Monoclonal antibodies have provided insight into the inhibitory mechanism of serpins[238], the regulation of stability against conformational change in PAI-1[185,186], structural differences between AAT polymers[69,76], the kinetics of intermediate formation[181] and allosteric regulation of stability[117]. From the screen, 1D9 was identified as having a preference to

monomeric AAT over polymeric forms. This is the first time that a monomer-selective antibody has been developed and potentially provides a direct complement to the polymer-specific 2C1 mAb.

4.4.2.1 The 1D9 epitope is most exposed in the native AAT monomer

The 1D9 mAb displayed a 10-fold preference for monomeric AAT over most polymer types of AAT. It also had an intermediate recognition of RCL-cleaved AAT. This was suggestive of a site of binding that was variant between the 5-stranded native monomer and 6-stranded cleaved AAT and most likely involved at least part of the RCL or β-sheet A strands 3 and 5. The reduced recognition of polymers compared to monomer was also in keeping with all the currently proposed polymerisation models, which agree that RCL insertion (regardless of being inter- or intra-molecular) is integral to the polymerisation mechanism[32,78,86].

Epitope mapping with the CPM fluorescent probe identified residues 360 and 283, out of the 25 AAT cysteine-mutants tested, to be part of the 1D9 epitope. These residues, in the RCL and s2C respectively, displayed a reduced conjugation efficiency to the probe in the presence of 1D9, indicating 1D9 was precluding labelling of these residues. This suggested that 1D9 was binding at the top end of AAT, in the vicinity of the so-called RCL 'distal hinge' region in a location distinct from the epitopes of other characterised AAT antibodies.

The crystal structure of the rAAT-1D9 Fab complex unambiguously defined the interaction between the molecules. It confirmed the epitope mapping findings and determined that the epitope is discontinuous (also referred to as conformational). 1D9 binds to the top end of the AAT molecule and interacts with residues on the C-terminal end of the RCL, s1C, s2C, s1B and s2B. This region of the molecule is suggested to be mobile, as it moves during the transition to the latent conformation and polymerisation[23,73]. Residues 283 and 360 identified from the CPM assay were involved in van der Waals interactions and neighbouring residues formed hydrogen bonds with 1D9. From the crystal structure, no other residues tested in the assay would be predicted to reduce conjugation efficiency, indicating the robustness of the epitope mapping technique. Interestingly, residue 346 at the N-terminal end of the RCL became slightly more exposed to CPM fluorophore with 1D9 binding. An overlay of native monomer (PDB ID:1QLP[8]) and the complex structure offers a possible

explanation of this observation; 346 in the complex appears more solvent-exposed, away from the body of the AAT structure (Figure 4.13). These two observations combined suggest that in solution, 1D9 binding may have a subtle effect on RCL conformation.

The structure of the 1D9 Fab and AAT complex is the first deposited crystal structure of AAT bound to an antibody. In fact, the 1D9 Fab alone is also the first structure of an antibody fragment specific for AAT. Both structures were equally important for model building and the structural analysis of the binding interaction between 1D9 and AAT. The high resolution of the Fab alone, together with sequencing results, meant that regions of sequence ambiguity in the antibody variable regions could be determined. Furthermore, as the complex was relatively low-resolution for X-ray crystallography, the Fab served as a reliable template in the complex structure. There are a large number of Fab structures in the PDB, such as the 1MF2 structure used during the structural determination of 1D9, which display the well-described β-sandwich immunoglobulin fold seen in the 1D9 structure[218].

The complex data were a little problematic in that there were large gaps in the electron density. This was particularly true for more peripheral regions of the AAT structure. It is unknown why there were such gaps given the good statistics of the data. It is possible that the high packing density of the crystal environment caused common crystal defects to lead to structural perturbations in different AAT molecules and this would explain the contrasting 50% difference in average B-factors between the two complexes. The crystal environment has been demonstrated to influence the structure of many other proteins[239]. Nevertheless, the complex structure was good enough to reveal the 1D9 epitope and the structural changes that occur on binding in both AAT and 1D9.

Being the most variable region in sequence, the complementarity-determining regions (CDRs) of antibodies form the antigen binding site. All but the L2 CDR of 1D9 formed contacts with the AAT molecule. Therefore, it is not surprising that the CDRs were the areas of largest conformational change (according to RMSD) and showed the greatest reduction in B-factor between the 1D9 free and bound states. The H3 CDR is often noted as being most important for antigen interactions and in this case, it made the most number of interactions with AAT, along with H2[168]. A change in the relative orientation between the variable and constant domains was

apparent between the complex structures in the asymmetric unit. Aside from that, other structural changes that have been observed in other antigen-antibody interactions, such as in the first loop of the CH1 domain, were not apparent[168]. Therefore, antigen binding does not appear to be associated with conformational changes in 1D9 far away from the antigen binding site[168].

The AAT structure also showed some conformational changes when comparing native and 1D9-bound structures. Predictably, the RCL that interacts with 1D9 was in an altered conformation between the two states. However, there were also structural changes in helix F and helix D which are at sites distant from the epitope. Conformational changes at sites involved in and distinct from the binding interface have also been observed in antigens upon complex formation[229,235]. It may be that these movements are the result of 1D9 binding or are simply in varied conformational states. There is an allosteric connection between the 'breach region' at the top of β-sheet A and helix D in heparin-binding serpins such as antithrombin and the hormone-binding protein corticosteroid binding globulin[3,234,240]. These helices are also in variant positions in crystal structures of serpins and structural plasticity of helix F has been observed in molecular dynamics simulations and biochemical studies[28,231,232].

The apparent contradictory results from the western blot and peptide-blocking ELISA, regarding the nature of the 1D9 epitope can also be explained from the complex crystal structure. Most of the AAT residues involved in interactions with 1D9 are in the RCL and s1C, residues which are continuous in sequence. Loss of some of the peripheral interactions in s2c, s1B and s2B in a denatured state may have an impact on affinity but provided a sufficiently high concentration is present in an assay, this difference may not be observed. The importance of the RCL/s1C interaction and their encapsulation in the 1D9 epitope is highlighted by the lack of recognition of cleaved AAT by 1D9 in the denaturing western blot: in the denatured state, the residues after the GluC cleavage site (354), are lost from the main body of AAT and hence cleaved AAT is not detected. The failure to prevent recognition of monomeric AAT by 1D9 that had been pre-incubated with a peptide corresponding to the C-terminus (359-394) at significantly higher molar ratios, probably reflects an increased avidity from the additional interactions outside of strand 1C. This indicates that these interactions observed in the crystal structure are also an important component of the binding interaction. The strong affinity that the 1D9 mAb has for AAT (2-12nM as measured

by SPR) is likely to out-compete weaker binding to a peptide that only partly constitutes the epitope. The affinity for 1D9 mAb and the peptides could be tested by SPR to assess the contribution of this linear part of the epitope, with the peptides conjugated to the chip and the antibody as the analyte.

Similarly, the affinity of 1D9 mAb for cleaved AAT could be measured by SPR. The anticipated higher binding constant can be structurally reasoned from knowledge of the 1D9 epitope. The protease GluC cleaves in a non-canonical position, 4 residues N-terminal to the site targeted by neutrophil elastase. Therefore, despite all of the residues identified as components of the epitope remaining in place, it is striking that 1D9 has a lower affinity for this form. One key difference is that the RCL is no longer constrained in the so-called 'canonical' conformation[8]. This in turn would suggest that the increased flexibility and lability of the RCL/s1C region disfavours binding, possibly by increasing the entropic cost of conformational selection [241].

The head-to-tail nature of polymers is likely to obstruct 1D9 binding to subunits, reducing the affinity of 1D9 to polymers. The major component of the 1D9 epitope - the distal hinge of the RCL - in the three polymerisation models is predicted to be perturbed relative to the native monomer but the known flexibility of the inter-subunit linkage may not prevent binding. For the C-terminal model, the epitope is similar to that for cleaved AAT except the epitope is comprised from residues in adjacent subunits (e.g. 360 from subunit n and 283 from subunit n+1). If this model was representative of *in vivo* polymers, the reduced affinity for polymers compared to cleaved AAT could potentially be reasoned by obstruction of 1D9 binding by neighbouring subunits (Figure 4.29). For both the loop-β sheet A and β-hairpin models, the epitope is formed from residues in one subunit. However, the epitope would be predicted to be significantly disrupted in the loop-β sheet A model of polymerisation due to the insertion of RCL residues into the neighbouring subunit. In the β-hairpin model, the extrusion of s5a into the neighbouring subunit would clash in the position that 1D9 Fab binds. Both these alternatives, would predict a lower affinity of 1D9 for polymers compared to monomers.

The 2C1 antibody has been able to shed some light onto aspects of the polymerisation mechanism[69,76,191] and it was considered whether the specificity of 1D9 could also provide such information. The epitope indicates that the most likely point of interaction with a polymer would be at the donor end. If the loop-β sheet A

mechanism is representative of the polymers considered, then each polymer would present a single native-like RCL which should be recognised well by the antibody, and in the β-hairpin model a similar region would be available as part of a more extensive extruded pair of β-strands. In contrast, the C-terminal model presents an unbound region, reminiscent of the peptide that failed to block antigen recognition, and would likely be recognised similarly poorly.

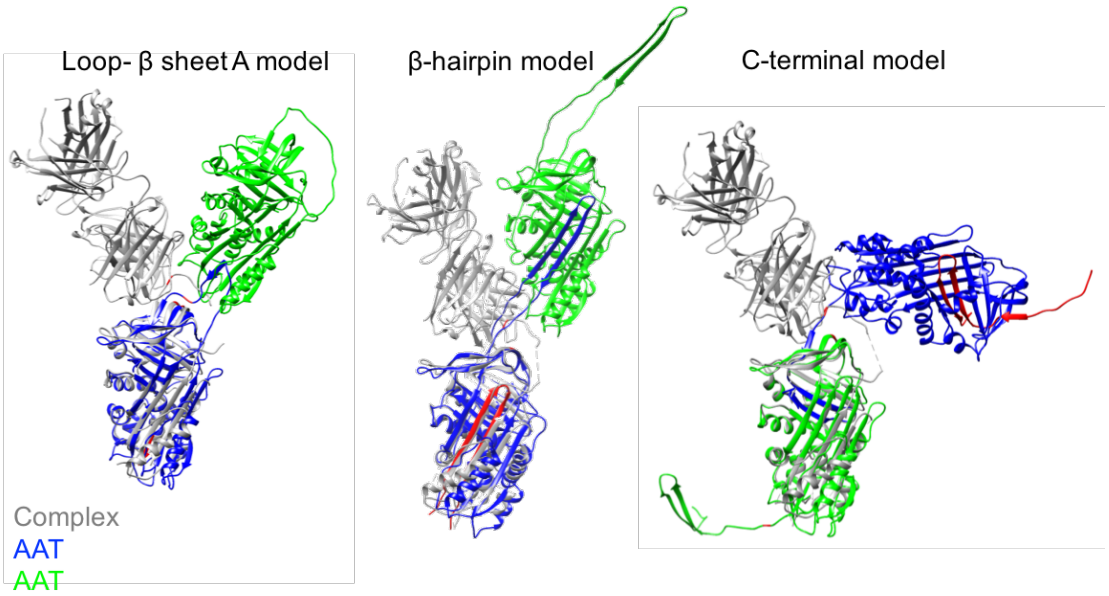


Figure 4.29. Superposition of the 1D9-AAT complex structure with models of the three proposed polymerisation mechanisms. *The crystal structure of the 1D9-AAT complex (grey) was superposed onto one AAT subunit (green and blue) for each polymerisation model. The residues 360 and 283 in the epitope region are highlighted in red.*

4.4.2.2 The RCL and s1C are important for polymerisation and inhibitory activity of AAT

1D9 is the second mAb that has been found to exhibit polymerisation blocking activity *in vitro*. Thermal shift experiments showed that the binding of 1D9 to AAT stabilised the native state from progressing to the intermediate during heat-induced polymerisation. Considering the 1D9 epitope, this activity is not surprising. Insertion of the RCL into β-sheet A is a common feature between the polymerisation models. Release of s1C in polymerisation has also been documented and accordingly heat-induced polymerisation did not occur when s1C was disulphide-bonded to s2C[73,74]. Therefore, 1D9 binding is likely to stabilise these monomer-specific structural

elements, stop conformational changes and prevent RCL insertion. This mechanism of action is in contrast to a mAb against transthyretin, which inhibits fibril formation by binding to the misfolded intermediate, rather than stabilising the native state[242]. Experiments with 1D9 were conducted for polymerisation induced by heating; it would be interesting to investigate whether 1D9 also blocked polymerisation induced by acid or denaturant. This would indicate whether the RCL and s1C are also important features of these polymerisation mechanisms. Likewise, 1D9 activity has not been tested on polymerisation of different disease variants of AAT. However, as mutations are believed to have similar structural effects, it is likely the action of 1D9 is translatable[65,82].

The mode of action of 1D9 is quite different to that of the polymer-blocking mAb 4B12. 4B12 lacks conformational selectivity and binds to helices A and I on the opposite side of the AAT molecule to 1D9, an area which is relatively invariant between AAT conformations[116,117]. This area was not previously associated with conformational modulation but 4B12 binding is believed to prematurely open β-sheet A, antagonising polymerisation as a result. Therefore, antibodies binding at very different sites on AAT can cause the same outcome, highlighting the interrelationship between structural elements of the AAT fold. When 1D9 and 4B12 were directly compared using an ELISA to monitor polymerisation of the Z variant under milder conditions than considered in the FRET experiments, 1D9 was shown to be the more efficient polymerisation blocking mAb of the two (personal communication, Dr E. Miranda and J. Perez). The centrality of the RCL to the polymerisation mechanism offers an explanation of this. An engineered single-chain variant of 1D9 co-transfected with Z-AAT in a cellular model could provide information on the conformational state of the RCL prior to polymerisation, similar to that carried out with 4B12[116,117].

RCL insertion into β-sheet A is key to AAT's physiological function of inhibiting proteases. Therefore, it was of no surprise when 1D9 binding efficiently stopped AAT's ability to inhibit chymotrypsin activity. The stoichiometry of inhibition (SI) for chymotrypsin was reduced to 23 in the presence of 1D9. This is the first time that a mAb has nearly eliminated the inhibitory activity of AAT. Despite 4B12 and 1D9 both being able to block polymerisation, 4B12-bound AAT nearly completely retained inhibitory activity against neutrophil elastase and trypsin[116,117]. MAbs against other serpins have however been developed that target serpin anti-proteolytic activity[186,238].

These antibodies have employed three different mechanisms: sterically blocking protease binding to the RCL, converting the serpin into an inactive conformation and inducing serpin substrate-like behaviour[186,238,243]. Based on the 1D9 epitope and superposition of the proteases with the 1D9-AAT complex (Figure 4.30), 1D9 is proposed to act through a mechanism of steric interference. This could be tested by determining association rate constants between AAT and protease in the presence of different amounts of 1D9. Antibodies against PAI-1 have bound in a similar location, with epitopes comprising the distal hinge of the RCL[188]. Similarly to 1D9, they react preferentially with the native rather than cleaved conformation and are proposed to interfere with formation of the protease-serpin complex[188,243]. An antibody against PAI-1 has also bound the N-terminal end of the RCL, in a region removed from the protease cleavage site and therefore distinct from that of 1D9[244].

A more unexpected result was the drastic difference of the anti-proteolytic activity of 1D9-bound AAT against neutrophil elastase compared to chymotrypsin. The SI compared to unbound AAT was still reduced but only to 1.8. It is worth noting that antibody binding is a reversible, equilibrium process, whereas the serpin mechanism of inhibition is an irreversible process. By definition, given sufficient time, the latter will always win over the former. The superposition of neutrophil elastase on trypsin from the trypsin-AAT complex structure (PDB ID: 1OPH[22]) with the 1D9-AAT complex structure appears to show a similar level of steric clash between 1D9 Fab and chymotrypsin (Figure 4.30). However, multiple studies have highlighted the sensitivity of the inhibitory mechanism and SI to single point mutations, changes in RCL length and protease identity[207-209,245]. For example, mutations in s1C have been documented to have opposite effects on the inhibitory activity of AAT: mutation of P10' did not alter activity towards thrombin but mutations at P6'-P8' caused the serpin to become a substrate rather than inhibitor[246]. Decreased inhibitory activity is reflected by a natural variant of antithrombin causing the deficiency associated with thrombosis[247]. Similarly, the importance of the length of the RCL and consequently, the rate of RCL insertion has also been demonstrated[207,208]. Lastly, variations in the SI have been observed when different proteases are inhibited by the same serpin[207,208,245]. This is thought to be due to the differential rates of deacylation associated with protease activity, which affects the balance between substrate and inhibitory behaviour of serpins[209]. Therefore, subtle changes in the interaction between AAT and chymotrypsin or neutrophil elastase due to 1D9 could influence inhibitory activity.

It is worth noting that the effect of 1D9 on the inhibitory activity of AAT against trypsin has not been tested yet and that the inhibitory activity of 4B12-bound AAT cannot be tested against chymotrypsin. Preliminary experiments showed that 4B12 itself reproducibly inhibited chymotrypsin (personal communication, Dr J. Irving). Therefore, it is unknown whether 4B12 does not affect AAT inhibitory activity against all proteases.

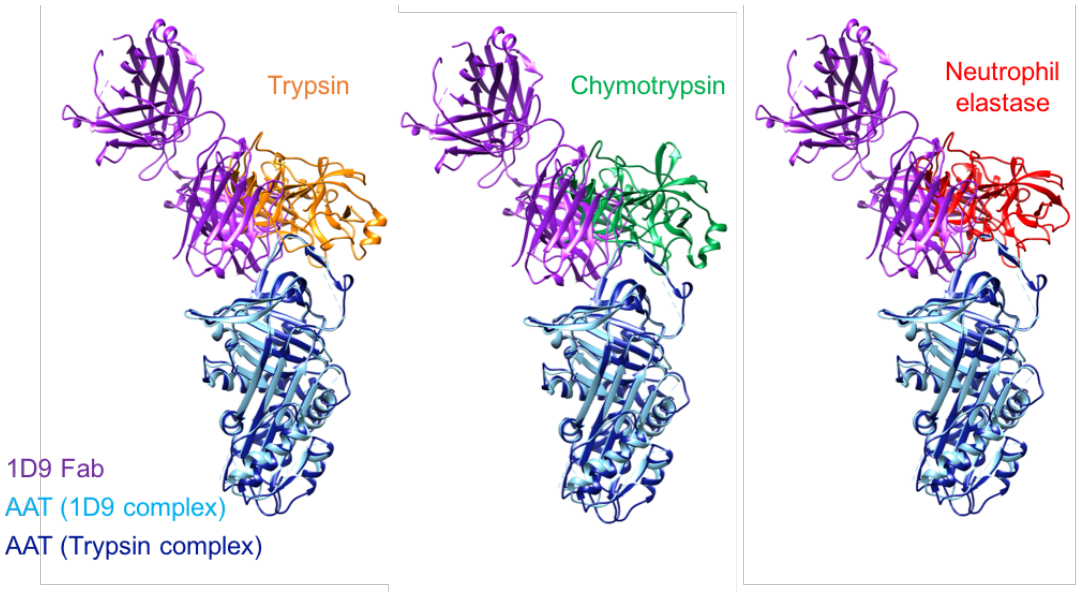


Figure 4.30. Steric clashes between the 1D9-AAT complex and protease. *The 1D9-AAT crystal structure (purple and light blue respectively) is superposed with the trypsin-AAT structure (PDB ID: 1OPH[22], orange and dark blue respectively). Chymotrypsin (green, PDB ID: 1CA0[248]) and neutrophil elastase (red, PDB ID: 3Q76[249]) were superposed on the trypsin structure using Chimera.*

4.4.2.3 1D9 provides the first evidence that liver and heat-induced polymers are distinct

A further interesting characteristic regarding the reactivity of 1D9 with different AAT species was the lack of recognition of heat-induced polymers in comparison to the other AAT polymers tested. Particularly noteworthy, is the large difference between the IC_{50}s of the plasma ZZ-AAT and liver ZZ-AAT polymers (151 and 168 ng/ml respectively) to the M heat polymers (> 400 ng/ml). Based on 2C1 reactivity, these polymer types have always been regarded as being similar in structure[76]. From the

2C1 epitope mapping, it is known a region of structural similarity is helices E and F[191]. Yet, 1D9 provides the first evidence that a different region of the polymer is structurally variant between the two forms. The 1D9 epitope mapping localises this difference to the RCL/s1C area of AAT which has implications for the polymerisation mechanism between heat-induced and *ex vivo* polymers. It has been proposed that the heat-induced polymer adopts the C-terminal configuration, although this study is complicated by use of a stabilising disulphide and the use of denaturant[86]. This form would not be expected to have an intact full 1D9 epitope at the terminus, while the loop-β sheet A model, and to some extent the β-hairpin model, would. However, an inter-subunit interaction would be more likely to occur with the C-terminal and β-hairpin models due to their increased flexibility.

In an attempt to use the 1D9 Fab to visualise the origins of the structural difference, liver ZZ-AAT polymers with and without prior 1D9 incubation were subjected to NS-EM. Unlike labelling of polymers with the 4B12 Fab (Chapter 5 and 6), no uniform labelling of polymers (identified by protrusions from the AAT chain) was observed with 1D9 Fab. The micrographs with and without 1D9 labelling could not be distinguished, indicating a low level or lack of labelling, and so the equivalent experiment for heat polymers was not carried out. The low labelling could be for several reasons. The 1D9 Fab may have a low affinity for polymers (the affinity constant for polymers was not measured by SPR) and at the dilutions used for NS-EM dissociated. A potential solution could be to crosslink the Fab to polymers, yet concentrations of crosslinker must be optimised to limit crosslinking of polymer chains and extensive aggregation. An alternative is that there may indeed be labelling, yet the size of the Fab and lack of repetitive labelling on each subunit may be too difficult to identify by NS-EM. Indeed, there are occasional protrusions that resemble Fab binding (Figure 4.28). This could be tested using a gold-labelled 1D9 Fab to provide an easily-identifiable intense spot of electron density, similar to what was done to verify polymer identity in early EM studies[41,81]. Lastly, the pH sensitivity of the 1D9 Fab-AAT interaction has not been investigated and the acidic pH (4.5) of uranyl acetate may dissociate the complex. A neutral negative stain such as phosphotungstic acid could be tried.

4.4.3 Implications of 1D9 characteristics

A therapeutic that can prevent polymerisation of AAT is desired for the treatment of AATD. Although 1D9 shows the desired activity, much like 4B12, its use as a therapeutic is limited due to the inaccessibility of antibodies to the ER of hepatocytes. The most promising and useful characteristic of 1D9 is through its monomer-selective reactivity. Currently, there is not a method to directly measure a patient's level of active AAT in the clinic. Instead, the total AAT concentration, irrespective of conformer, is calculated by nephelometry using a polyclonal antibody. Based on correspondence with the product support team of a major clinical assay provider, it appears that the influence of the presence of polymers on the value obtained has not been considered or established (J. Irving, personal communication). Therefore, 1D9 could have application as a diagnostic reagent. To test its suitability, plasma ZZ-AAT samples will be quantified for different AAT conformers using ELISA with 1D9, 2C1 and 3C11 as detection antibodies.

5.0 Negative stain EM (NS-EM) of heat polymers

5.1 Introduction

5.1.1 Electron microscopy of serpins

The structure of the AAT monomer and other serpins in varying conformations have been solved to high resolution by X-ray crystallography[8,49,82], including as circular polymers and in dimeric configurations that appear to be the result of contacts formed within the crystal lattice[5,78,84-86]. These latter structures have been generated using recombinant material, polymerised under artificial conditions, and lacking complex glycosylation states. Linear AAT polymers, particularly those obtained from human tissue, have characteristics that make them unsuitable for X-ray crystallography or NMR: large size, flexibility, glycosylation and heterogeneity in length. Electron microscopy (EM) and modern single-particle reconstruction approaches provide a means by which structural information can be obtained from such heterogeneous samples.

NS-EM has previously been used to visualise polymers of AAT, anti-chymotrypsin and neuroserpin, which show a beads-on-a-string appearance[32,41,81,94]. Polymers have also been immunolabelled with colloidal gold-conjugated antibodies confirming their identity[41,81]. There are no published 3D reconstructions of polymers from EM data but prior work in the lab has yielded 3D reconstructions of a dimeric repeating subunit in polymers purified from patient tissue by NS-EM. These were of low resolution (~30Å) and cryo-EM of this material has the potential to obtain higher resolution envelopes (Chapter 6). High molecular weight, symmetrical and rigid molecules generate the highest contrast in EM images and facilitate alignment of particles. However, the small size of the AAT subunit, combined with flexibility in the polymer and lack of symmetry or strong features produces low signal-to-noise images, and presents significant challenges to image processing.

5.1.2 Previous negative stain-electron microscopy (NS-EM) image analysis of the *ex vivo* polymer

In a prior study, polymers were purified from the liver of a Z-AAT homozygote and decorated with the Fab domain of the 4B12 mAb[250]. This antibody was selected due

to its lack of conformational preference (which might introduce bias as to the components of the sample that are bound), and biochemical data indicating the site of the epitope[116,117]. The antibody fragments increased the size of the subunits, aiding particle picking and also helping to determine the orientation of particles during 3D reconstruction. NS-EM using 2% w/v uranyl acetate showed polymers were mainly linear, with a few being circular, in agreement with early EM studies.

This NS-EM dataset was processed using the IMAGIC image processing software[121]. A total of 8120 dimer particles were manually selected from longer polymers for single particle analysis and aligned to one monomer. The dimer is the smallest polymer subunit and so reduces – but does not eliminate – the impact of the flexibility of the inter-subunit linkage with respect to longer polymers. Eigenimage analysis indicated that the images were better divided into two subsets, a 'compact' and an 'elongated' dimer of 54% and 46% frequency, respectively. Using a starting model based on duplication of the cleaved crystal structure of AAT (used as a surrogate for a hyperstable, 6-stranded β-sheet A subunit conformation) and information on the epitope of 4B12, initial models were generated to seed 3D reconstructions for each subset. After multiple rounds of refinement, the resulting envelopes were strongly resolution-limited to around 30Å, though still permitted an assessment of the compatibility of different models of polymerisation due to the asymmetry introduced by the bound Fab molecules. The calculated distance between the centres of mass of the subunits for the compact subset was 68Å with a rotation of 16° between them. For the elongated subset, this was 73Å and 11°, respectively (Figure 5.1). It is not known whether these distinct forms are interspersed within polymer chains or between chains. Furthermore, the significance of these two populations *in vivo* is unclear; it is possible they represent two extremes of a continuum of conformations, given the flexibility inherent in polymers.

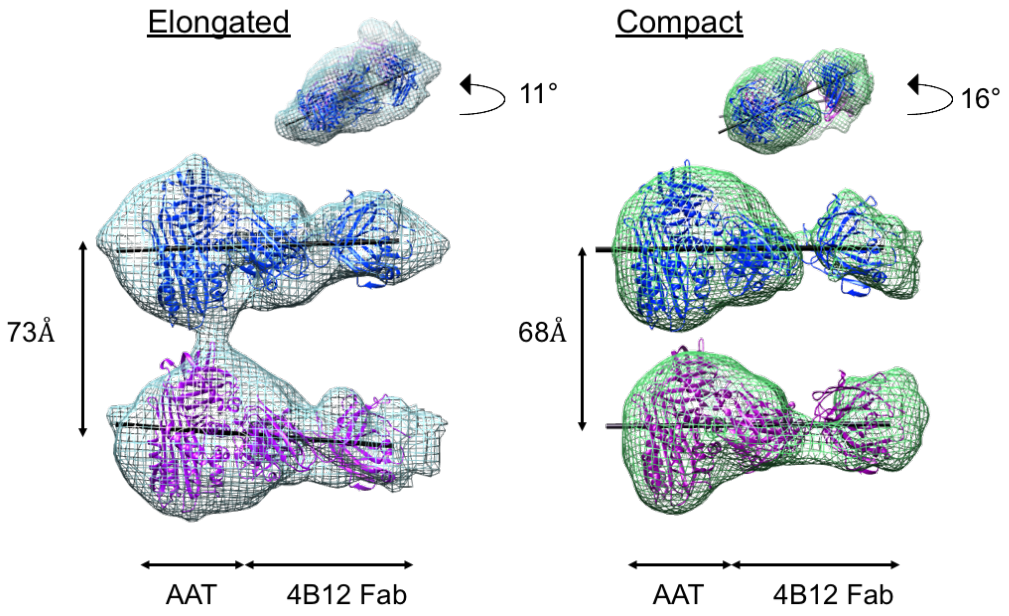


Figure 5.1. 3D reconstructions of the AAT dimer isolated from ex vivo material and labelled with 4B12 Fab. *Two subsets were identified, termed elongated (left) and compact (right) and a model of cleaved AAT in complex with 4B12 Fab was fitted into each. Axes were drawn through the middle of the complexes and used to determine the distances between subunit centres. Work carried out by Dr S. Faull, Dr J. Irving and Prof E. Orlova and figure reproduced with permission*[250]*.*

The three major proposed polymer models were fitted into the envelopes determined from the NS-EM using the PHENIX software package[221]. As the envelopes are low-resolution, the fitting could only be informative in identifying models that are inconsistent with the data. This analysis showed that the loop-β sheet A model was poorly accommodated within the densities of both subsets, the β-hairpin could be modelled into the compact dimer density but not easily in the elongated density and the C-terminal model could be modelled into both subsets. Therefore, the C-terminal model was the most favoured out of the three by the NS-EM data. This work represented the first time that 3D reconstructions of *ex vivo* AAT polymers from EM data had been performed.

5.1.3 Assessing the relevance of *in vitro* artificial polymers to those found *in vivo*

Most research into the mechanism of AAT polymerisation has used artificial means to generate polymers. This is achieved by heating or incubating monomeric AAT with acid or denaturant[47,66,68,69]. It is believed that different polymeric forms are generated by the different methods and this is what underlies the proposal of three different inter-subunit linkages in the major polymerisation models[32,50,69,78,86]. There have been disagreements in the literature about which method of artificial polymerisation represents that which occurs *in vivo*. However, the common recognition of heat-induced polymers and *ex vivo* polymers by the mAb 2C1 favours heating as a suitable artificial method of polymer formation[69,70,76]. Polymers formed by Gu-HCl or acid incubation are not recognised by 2C1 and therefore the β-hairpin model, based on the crystal structure of an antithrombin dimer formed at low pH, is believed to be unrepresentative of *in vivo* polymerisation[69]. Understanding whether heat-induced polymers are formed by the same linkage as *ex vivo* polymers would greatly aid laboratory research.

The aim of this chapter was to obtain NS-EM 3D reconstructions of heat-induced AAT polymer and structurally compare the two types of polymers. Image processing of the data was also compared for two different software packages: IMAGIC[121] and Relion[122].

5.2 Methods

Sample preparation, NS-EM grid preparation and NS-EM imaging are covered in method 3.5.4.

5.2.1 NS-EM data collection of the heat-induced polymers

The heat polymer/4B12 Fab dataset was collected on a Tecnai 200 kV FEG with a Direct Electron (DE-20) direct detector. A dataset of 169 micrographs was recorded at 29,000X nominal magnification with a defocus range of 0.2-1µm. The pixel size for images was 1.54Å and movies with 18 subframes (1.8s total exposure) were collected and aligned using EMAN2[251] and IMOD[252].

5.2.2 Image processing of the heat-induced polymer using IMAGIC

5.2.2.1 CTF correction and particle picking

The images resulting from the movie subframe alignment were corrected for the contrast transfer function (CTF) using CTFFIND3[253]. CTF correction and astigmatism was manually checked using the Fourier transform of images. No micrographs were discarded. CTF-corrected images were visualised in EMAN1 Boxer and a 5x5 median filter applied to aid selection of polymer particles[254]. Helix mode – a semi-automated process in which the start and end of linear sections of polymers were manually selected – was used, where particles were picked into a 200x200 pixel box, and neighbouring boxes had a 140-pixel overlap. Box coordinates were saved for each micrograph and applied to the unfiltered CTF-corrected images to create the dataset of a total of 17,994 boxed particles (AAT/Fab tetramers) for single particle analysis processing in IMAGIC.

5.2.2.2 Alignment of particles

Particle images were normalised to 1σ using a circular variance mask and band-pass filtered to remove high-frequency background. They were rotationally and translationally aligned to a 55x200 pixel low-pass filtered vertical rectangle. They were then classified and averaged by multivariate statistical analysis (MSA). The 10 “best” class averages that represented different views of the tetramer particles and had well delineated polymer subunits were used as references in multi-reference

alignment. The images after the second alignment were re-subjected to MSA. Eigenimages were used to reveal the major variations in the dataset. Images were classified according to selected eigenimages with visually interpretable features, for example those that showed vertical misalignment or horizontal misalignment. The subsequent alignment and classification round then focused on the corresponding translation and was against class average references or the total sum. Improvements in the alignment were apparent in the increasing definition of subunits in the total sum of individual classes. Permitted translations were decreased gradually at each iteration from 40 pixels to 5 pixels. At this stage, there were four subunits per image.

After six rounds of alignment the aligned tetramer images were trimmed to focus on the central dimer. The total sum of all particles guided the boundaries that were selected to yield the dimer: the tetramer box length was cut off at 45 pixels to 145 pixels. The dimer was then centred in a 200x200 pixel box and a soft-edged 100x180 pixel rectangular mask applied to focus alignment on the dimer. The alignment of the dimer was improved in a similar manner to the tetramer, using MSA and classification based on selected eigenimages that reflected the dominant basis for the variability in the subset. “Dimer” particles that only had one AAT/Fab subunit in the particle box, or that were poorly defined, were removed after sorting by MSA, resulting in 12,823 particles.

5.2.2.3 Division of the dimer particles into “small-angle” and “big-angle” subsets

The remaining 12,823 dimer particles were classified into class averages with 25 particles per class, based on 6 active eigenimages. They showed that there were differences between the images in respect to the angle between the subunits. The class averages were divided into two subsets, the small-angle and the big-angle subsets, by visual inspection of the position of the Fab fragments. The class averages were low-pass filtered to aid selection. Those that could not be categorised initially in this way were selected for re-classification. The raw particles comprising all the selected classes were extracted resulting in 6,036 particles in the small-angle subset and 2,643 in the big-angle subset.

5.2.2.4 Projection matching for the small-angle subset

The elongated 11° dimer model generated from the prior *ex vivo* NS-EM dataset[250] was low-pass filtered to 30Å in IMAGIC and used as a starting model for projection matching to the small-angle subset images. This involved projection at 10° increments around the γ axis between 0-360° and between 70-110° around β to produce 180 2D references that were normalised to 1σ. The α angle was kept at 0° (the α, β and γ angles referred to are according to the IMAGIC Euler angle convention). The small-angle subset class averages (25 images per class) underwent projection matching to these references based on cross-correlation scores and thereby assigned Euler angles. The assigned projections were used to reconstruct an initial 3D model, and a subsequent 3D model was created based on the top 50% of class averages according to a 3D-error score. The model was masked in IMAGIC, converted to the ".mrc" map format and assessed in Chimera. This model was then low-pass filtered, reprojected and matched in an iterative process. During this process, the number of images per class average was reduced to 10 and the reprojections were sampled at finer 5° intervals. After the angle between subunits stopped changing significantly, the process was stopped.

5.2.2.5 Analysis of the final 3D reconstruction for the small-angle subset

The final model was visualised in Chimera[9]. A threshold was applied to the volume at a value of 360Å^3: corresponding to 200 kDa (expected molecular weight of the AAT dimer bound to Fab), multiplied by 50% (a 'fudge-factor' taking account for negative stain and glycans) multiplied by 1.2 (an approximate value of Å^3/Da for globular proteins[255]). The atomic coordinates of a model of cleaved AAT bound with 4B12 Fab was fitted into each subunit in the dimer in head-to-tail orientation. The atomic models were also positioned in the opposite direction and the best fit was selected based on the cross-correlation score. An axis through the central mass of each AAT/Fab complex was used measure the distance and rotation between subunits.

The resolution was assessed in IMAGIC by randomly dividing the dataset into two groups, generating 3D reconstructions and comparing them by Fourier Shell Correlation. GraphPad prism was used to plot the FSC curve and fit a sigmoidal function. This was used to calculate the resolution at threshold values of 0.5 and 0.143, both of which are in common use[122,157].

5.2.2.6 Attempts at projection matching for the big-angle subset

Initial models for projection matching were created in Chimera[9]. Density for a single AAT/Fab subunit was created by excising, superimposing and averaging the individual subunits of the prior 11° *ex vivo* dimer. The average subunit was duplicated and translated 70Å based on the inter-subunit distance from measurements of the class averages. One subunit was rotated 90°-160° in 10° increments relative to the other, to generate 8 initial models at different orientations. These were then reprojected at 10° increments around the γ axis between 0-360° and between 80-100° around β to produce 108 2D references. The reprojections were used as references for competitive projection matching to the big-angle subset class averages. The number of matches to each set of references was computed and was greatest for models between 140-160° rotation. Projection matching to these models was undertaken and the resultant 3D reconstructions assessed in Chimera. No attempts yielded a model with two defined subunits, however.

5.2.3 Image processing of the heat-induced polymer using Relion

5.2.3.1 Particle picking and 2D classification

The motion-corrected and CTF-corrected micrographs used in the IMAGIC processing were imported into Relion v2.1[122]. An early 50Å low-pass filtered 3D reconstruction from the Relion *ex vivo* polymer data processing (detailed in 5.2.4.1) was used to generate 30 projections as references for autopicking. The parameters used for autopicking were a picking threshold of 0.05, minimum inter-particle distance of 60Å and maximum standard deviation noise of 1.5. Particles were extracted into a 150-pixel box, masked by a 190Å circle and the subset pruned over several rounds of 2D classification into 100 classes with manual subset selection. This left 25,323 particles.

5.2.3.2 3D classification

The particles were used as input into a 3D classification job using the Relion *ex vivo* polymer 60° model (detailed below) which had been filtered to 30Å as a reference. Five reconstructions were generated. The best reconstruction was subjected to a threshold at a volume corresponding to 1.5x molecular weight and 1.2Å^3/Da, filtered

using Chimera and used as a reference in another round of 3D classification against all 25,323 particles. Particles from the best reconstruction were selected (6,923 particles).

5.2.3.3 3D refinement and structural assessment

The 6,923 particles were refined against a 30Å low-pass filtered monomeric subunit reference (AAT/Fab) to generate the final reconstruction of the dimer. This was generated by erasing one of the subunits of the Relion *ex vivo* polymer 60° model in Chimera[9]. Model measurements was made as described in method 5.2.2.5. A mask was created in Relion based on the refined structure, to which a threshold and a Gaussian filter had been applied (with a standard deviation of 10Å) and a 20-pixel soft boundary added in Relion. This mask was used in the Relion post-process function for the estimation of the resolution using the FSC.

The 6,923 particles were subjected to 2D classification into 50 classes. A 3D *ab initio* reconstruction, based on the top 99% of the 6,923 particles when ordered by posterior probability computed during the 3D refine job, also generated a similar reconstruction to the final model. Further sub-classification of the 6,923 particles by 3D classification into 2 and 4 classes yielded structures with similar characteristics to the final model.

5.2.3.4 Further Relion jobs to identify other reconstructions

The 25,232 particles were used as input for a 3D classification job against the Relion *ex vivo* 88° model or the IMAGIC heat 16° model as a reference. Out of the 5 classes, images were selected that comprised the best-defined reconstructions. These images were used an input for a 'gold-standard' 3D refinement using the monomer subunit as a reference but no reasonable dimer reconstructions resulted. The 6,036 particles that had been selected during the IMAGIC processing also did not provide a consistent reconstruction when refined against a monomer reference.

5.2.4 Image re-processing of the *ex vivo* polymer using Relion

5.2.4.1 Particle picking and 2D classification

The motion-corrected and CTF-corrected micrographs that had been collected previously and used for processing in IMAGIC[250] were imported into Relion v2.1. The 11° *ex vivo* IMAGIC dimer was filtered to 50Å and reprojected at angles distributed around the Euler sphere using the Relion command line. Twelve references were generated for autopicking from the micrographs using a 0.025 picking threshold, 60Å minimum inter-particle distance and no maximum standard deviation of noise (-1). Only 3,512 particles were obtained and extracted into a 150-pixel box. They were masked with a 180Å circular mask and classified in 2D into 25 classes. Classes were selected with dimer particles, yielding 1,962 particles. These were used as input for a 3D classification job into 3 classes using the same 11° *ex vivo* IMAGIC dimer as a reference. One output poorly-defined model was low-pass filtered to 50Å. Particle picking and particle extraction was then carried out as in method 5.2.3.1. A total of 176,957 particles were extracted and pruned using several rounds of 2D classification into 100 classes with manual subset selection. A set of 68,528 particles and a more-refined set of 15,911 particles resulted.

5.2.4.2 3D classification

The more-refined set of particles (15,911) were used as input into a 3D classification job using the poorly-defined autopicking model as a reference. Four reconstructions were generated. The best reconstruction was thresholded at volume corresponding to 1.5x molecular weight (in this case molecular weight was judged to be 150kDa as one subunit had Fab density missing) in Chimera. It was then filtered to 40Å and used as a reference in the next round of 3D classification of the 68,523 particles to generate eight reconstructions. Classes were judged on visual appearance of the reconstructions, their constituent particles selected, and the process repeated. The particles were then classified twice in 3D against a single subunit reference – made by erasing the second subunit from one of the reconstructions in Chimera – to generate four reconstructions.

5.2.4.3 3D refinement and structural assessment

The 6,848 particles and 9,239 particles that corresponded to the well-defined models resulting from the 3D classification job were selected. They were refined and structurally assessed by *ab initio* reconstruction as in method 5.2.3.3 with the exception that no removal of particles based on the posterior probability score was undertaken.

5.2.4.4 Further Relion jobs to identify other reconstructions

The remaining unclassified particles (14,892) were used as input for a 3D classification job against the monomer or the IMAGIC *ex vivo* 11° model as a reference. No reasonable reconstructions were generated after image selection and refinement. The 8,120 particles that made up the IMAGIC *ex vivo* dimers also did not return a comparable dimer model when refined against a monomer reference.

5.2.5 Generation of an initial AAT/4B12 Fab complex

During assessment of the NS-EM reconstructions, a molecular model of the complex between AAT and the 4B12 Fab was used. This was based on the crystal structure of the cleaved monomeric AAT, after removal of the protease coordinates from PDB entry 1EZX[13]. Polymers isolated from liver have high-mannose glycans and those made from purified plasma protein contain mature glycosylation. Thus, glycan chains were added using the GlyProt server to residues 46, 83 and 247[256]. This AAT structure was chosen as the polymer models with the greatest supporting evidence all propose that the RCL is in the inserted conformation. The 4B12 Fab model had been produced previously using Rosetta Antibody modelling software based on the sequence of 4B12 variable regions. Modelling of the interaction between AAT and 4B12 Fab was carried out using PyRosetta software[164], which used the NS-EM single-subunit density, calculated as the average of both subunits in the NS-EM 60° liver reconstruction, and known interactions between residues in the 4B12 epitope (32, 43, 306, 261 and 266) and the paratope as constraints. One thousand atomic models were generated with randomly oriented subunits and were assessed by their fit into the EM density (expressed in Rosetta as a score roughly equivalent to the cross-correlation). The 20 best fitting models were analysed and the top models with AAT in opposite directions were chosen.

5.2.6 Fitting of the polymerisation models into the NS-EM dimer reconstructions

The three AAT polymerisation models had been generated for the previous NS-EM study[250]. Briefly, the C-terminal polymer model was prepared by rearranging the breaks between polypeptide chains in PDB entry 1D5S[80], which corresponds with a cleavage-induced polymer, resulting in a domain swap involving the C-terminal 35 amino acids. The β-hairpin model was prepared starting with the structure of cleaved antitrypsin (1EZX[13]), and unravelling and reconnecting strands s4A and s5A using Coot[161], with optimisation of geometry using PHENIX[221]. The loop-β sheet A model was based on the structure of intact AAT incorporating two peptides (of sequence FLAA) in the centre of β-sheet A[250]. Removal of the peptides and incorporation of the RCL was undertaken with Coot[161] and PyRosetta[164]. The best subunit models (329 and 787) were superimposed onto the linear polymerisation models to redefine the AAT/Fab interaction using Chimera. PyRosetta was used to randomly generate 500 starting models that differed in the relative positioning of subunits[164]. The AAT/Fab subunits were treated as rigid bodies and the linkers defined as flexible regions. The models were allowed to optimise into the reconstructions that were thresholded to a volume of 360Å^3 and were given a score based on the fit which was proportional to the cross-correlation and reflected the linkage-specific restraints. The fits of the models to the densities were evaluated for significance using the Dunn's non-parametric multiple comparisons test using GraphPad. To do this, the initial 329 or 787 subunit structures were superimposed over the two subunits in each top 10 scoring dimer, generating a starting CC value, using the superpose program in CCP4. Phenix real space refine was used to optimise the superimposed subunits into the density without the restraints imposed by the different polymer models and resulted in a resulting CC. The similarity between each model-restrained dimer and the final optimised model-free subunit positions was calculated as a fraction and the values averaged for each combination. A significance test was carried out assuming a common mechanism between liver polymers by combining the 60° and 88° fits. A second test was carried out assuming a common mechanism between liver polymers and heat polymers by combing the liver 60° and 88° and the heat 60° fits.

5.3 Results

5.3.1 Purification of AAT and production of heat polymers

Native AAT was purified using a two-step purification from plasma that had been rejected by the National Blood Service. The first step used the commercial Alpha$_1$-Antitrypsin Select resin (GE Healthcare) consisting of camelidae-derived single domain antibody fragments that are selective for AAT (Figure 5.2A, B). Elution was achieved by a high concentration of $MgCl_2$ and a denaturing gel showed large amounts of protein at a molecular weight consistent with AAT (52kDa). Fractions 3-9 were pooled and subjected to a Q-sepharose column to further purify AAT away from higher molecular weight contaminants (Figure 5.2C, D, E). A denaturing gel of the elution fractions showed pure monomeric AAT protein in earlier fractions with a possible covalent dimeric band appearing in later fractions. A non-denaturing gel subsequently showed fractions 16-18 to be predominately polymeric. Purified monomeric AAT was buffer exchanged into PBS and polymerisation induced by heating at 55°C and a concentration of 0.2 mg/ml for 48 hours. These conditions were sufficient to achieve near-complete polymerisation of the sample[74]. A distinct laddering pattern was seen on a non-denaturing gel for the heat polymer, whereas native polymer purified from plasma using an identical process showed a laddering pattern at a lower molecular weight but more of a smear at longer polymers (Figure 5.2F).

5.3.2 Imaging of heat polymers by NS-EM

Polymers of AAT formed by heating are the only *in vitro* induced polymer that shares an epitope with those extracted from hepatocytes, indicating at least a degree of structural similarity[76]. Polymers of Z-AAT isolated from hepatocytes and heat-induced polymers have been visualised previously by NS-EM and appear to have a qualitatively comparable ultrastructure[32]. However, a more detailed comparison between the two at the level of a 3D reconstruction has never been carried out.

The heat polymer, undecorated by Fab, was imaged by NS-EM. The heat polymers had the characteristic beads-on-a-string appearance (Figure 5.3), with a high degree of flexibility. Numerous small circular polymers were observed. Individual subunits could be distinguished within the polymer chain and appeared to have some but limited visual asymmetry.

It was deemed that obtaining a 3D reconstruction of the polymer by EM would be a substantial challenge as the AAT subunit is small and relatively featureless making the determination of angles of the 2D projections difficult. This is partly why an approach had previously been adopted where a molecule with a known binding site – the 4B12 Fab – has been added to *ex vivo* polymer samples prior to visualisation in EM experiments.

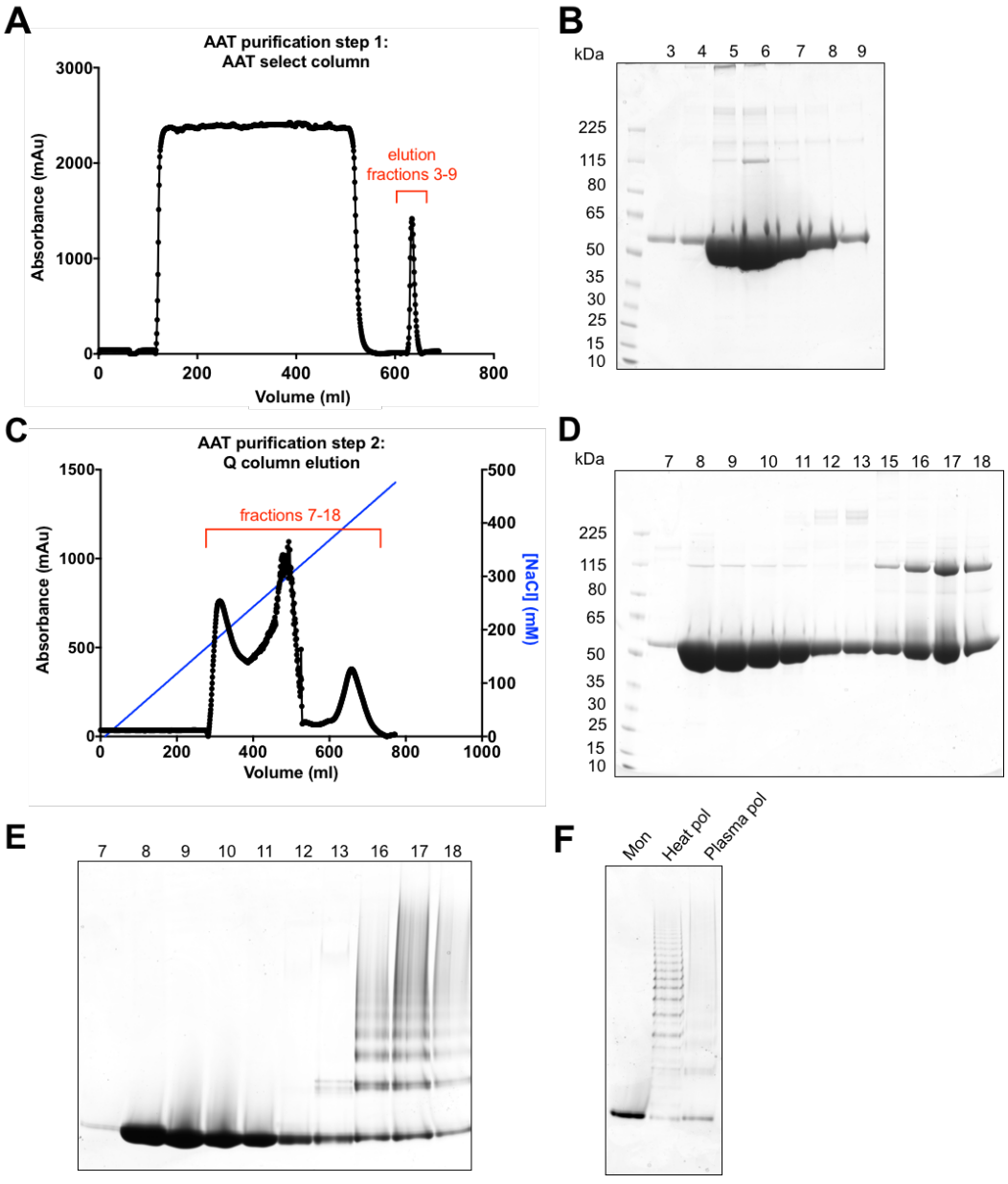


Figure 5.2. Purification of AAT from plasma. *A) A chromatogram showing a purification using the Alpha-1-Antitrypsin Select column. The high absorbance readings from ~100-500ml*

represent plasma proteins that do not bind to the column. Elution with $MgCl_2$ was started once a baseline had been reached following sample loading; a sharp peak is seen. B) A 4-12% w/v acrylamide reducing SDS-PAGE of elution fractions 3-9 showing a high concentration of protein at ~52kDa. C) The chromatogram for the Q-sepharose column. Three absorbance peaks can be seen, eluting from the column at approximately 200, 300 and 400mM NaCl concentrations. D) A 4-12% w/v acrylamide SDS-PAGE gel of fractions 7-18 showing the AAT band, as well a potential dimer band. E) A 3-12% acrylamide non-denaturing gel showing fractions 7-13 comprise monomeric protein whereas 16-18 contain polymers. F) A 3-12% acrylamide non-denaturing gel showing the characteristic laddering pattern of heat polymer.

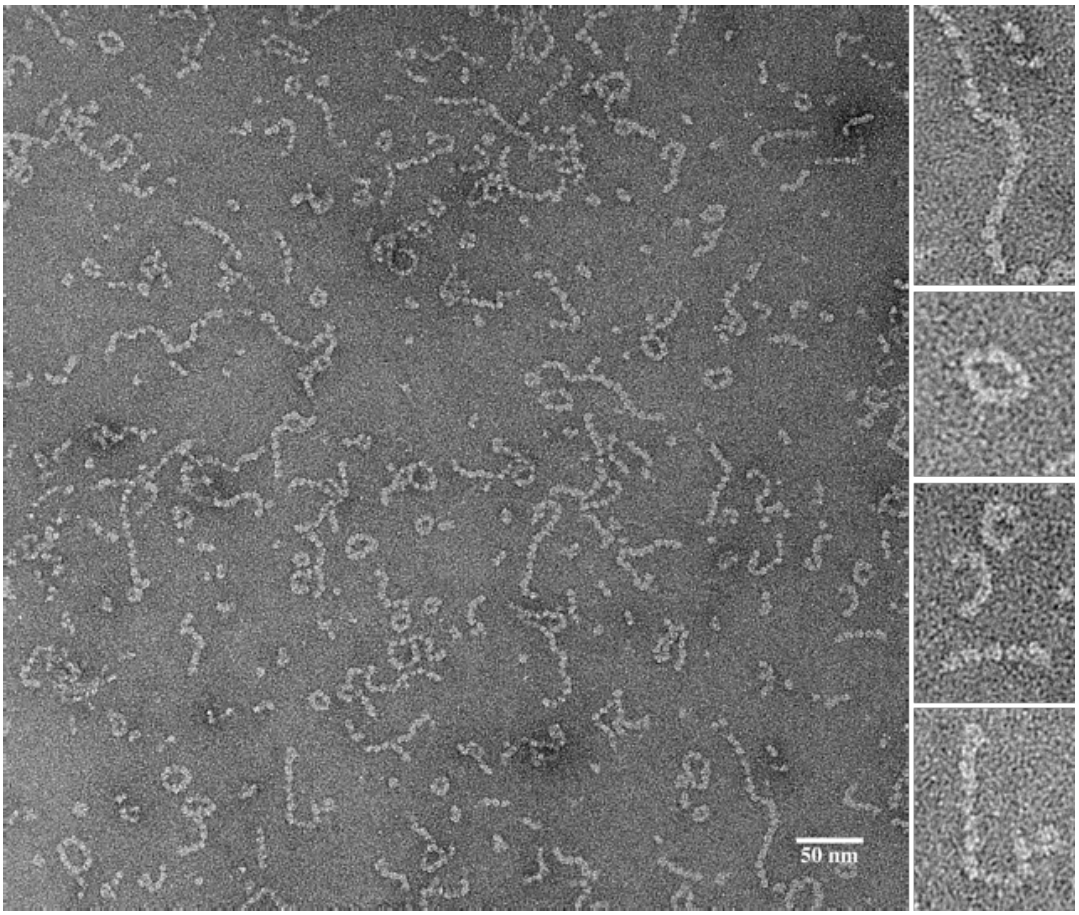


Figure 5.3. A negative stain micrograph of heat-induced AAT polymers. *The sample at 0.06 mg/ml was stained with 2% w/v uranyl acetate stain. The sample was visualised using a Tecnai T12 microscope equipped with a CCD camera at a nominal magnification of 67,000X. Scale bar, 50nm. Single polymer chains are shown on the right*[70].

5.3.3 Imaging of heat polymers bound with 4B12 Fab

Previous work undertaken by Dr S. Faull involved single particle analysis and 3D reconstructions of *ex vivo* polymers bound with 4B12 Fab[250]. *In vitro* heat polymers

are thought to represent a good model of *ex vivo* polymers due to the presence of a common epitope, as recognised by the 2C1 antibody[76]. However, this does not demonstrate structural similarity, it merely does not exclude heat-induced polymers as an appropriate model. This is particularly the case as a typical epitope covers only a small proportion of the total surface area (up to ~1,000 Å^2 or about ~5%)[227].

To allow a direct comparison between low-resolution structures, heat polymers were labelled with 4B12 Fab and imaged by NS-EM. The 4B12 Fab molecules are ~50kDa in size and were generated from the full-length mAb by ficin digestion and a protein A column purification step. Fab fragments do not bind to a protein A column, and thus are present in the flow-through whereas the Fc region (also ~50kDa) and undigested mAb (~150kDa) are retained (Figure 5.4).

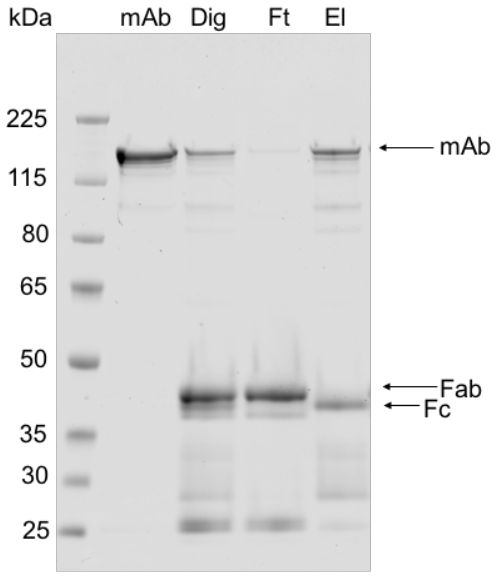


Figure 5.4. Generation of the 4B12 Fab. *An SDS-PAGE (non-reducing) gel showing in lane 1 the monoclonal antibody, in lane 2 the ficin digest, in lane 3 the flow-through from the Protein A column (Fab) and in lane 4 the elution from the Protein A column (Fc and undigested mAb).*

Binding of the 4B12 Fab to the polymers increased the size and changed the appearance of the particles (Figure 5.5). The Fabs could be distinguished from AAT on the majority of subunits and added visible protrusions from the polymer chain. These identifiable features should help the assignment of angles to the 2D projections that is required in the process of generating 3D reconstructions. Examples of the Fabs lined up on one side of the polymer, and where they appeared

on alternating sides, could be observed, highlighting the flexibility in the polymer chain and the linkage. Circular polymers were still observed. As polymers are extremely stable – they dissociate in around 6 M guanidine and have a melting temperature above 100°C – it is unlikely that the polymer linkage is modified due to the binding of the antibody, although its orientation on the grid in the presence of stain may be.

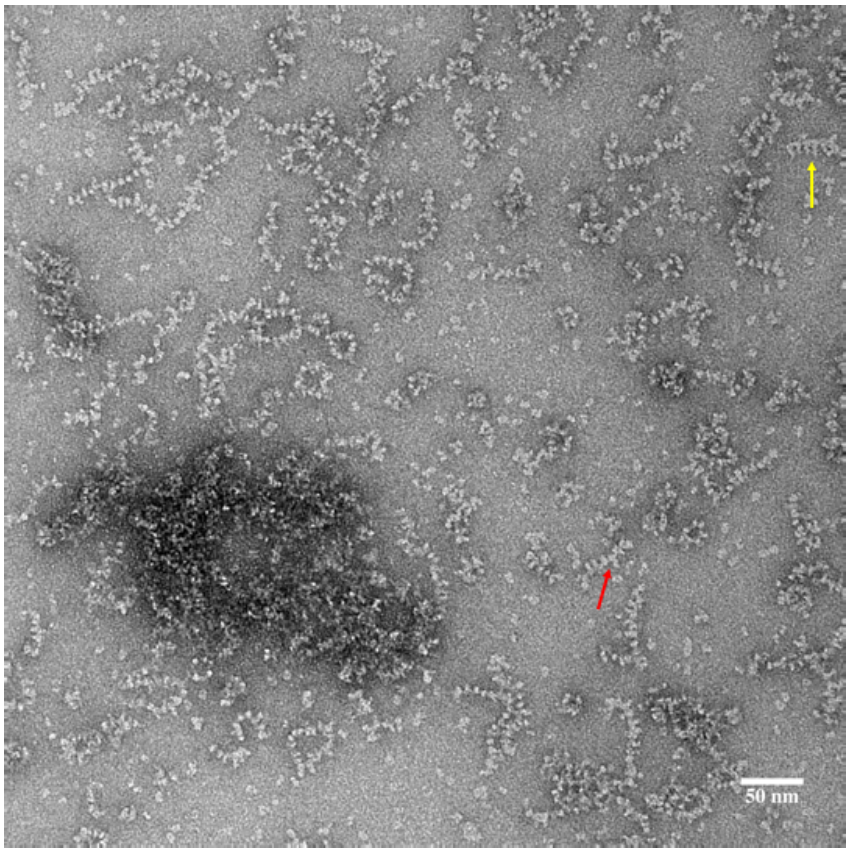


Figure 5.5. A negative stain micrograph of heat-induced AAT polymers bound with 4B12 Fab. *Sample at 0.07mg/ml was stained with 2% w/v uranyl acetate and visualised using a Tecnai T12 microscope at nominal 67,000X magnification. The micrograph was recorded at 1µm defocus. Yellow and red arrows highlight where Fab fragments are lined up on the same or opposite sides of the polymer, respectively. Scale bar, 50nm.*

5.3.4 Image processing of heat polymers bound with 4B12 by IMAGIC

IMAGIC is an image processing software that has been in development since the 1980s[121,162]. It uses a command-line format to run different programs (such as 2D alignment and 3D reconstruction) to efficiently process EM images. Image

processing is carried out in real space and it implements multivariate statistical analysis (MSA) and eigenvectors to simplify the analysis of image variations within large datasets. IMAGIC requires greater user-input and knowledge of image processing than more recent software, however it has flexibility and transparency in processing jobs.

5.3.4.1 Selecting AAT particles from micrographs

A NS-EM dataset of the heat polymer with bound 4B12 Fab was collected using a Tecnai 200 kV FEG microscope with a DE-20 detector. For single particle analysis, linear regions of polymer chain were selected using the helical picking function of EMAN1 Boxer (Figure 5.6)[254]. Just under 18,000 particles were manually picked from 169 CTF-corrected micrographs. The box size was chosen to be 200x200 pixels as this would accommodate the Fab fragment bound on either side of the AAT molecule. There was an overlap of 140 pixels between boxes. Examples of the boxed particles are seen in Figure 5.6B.

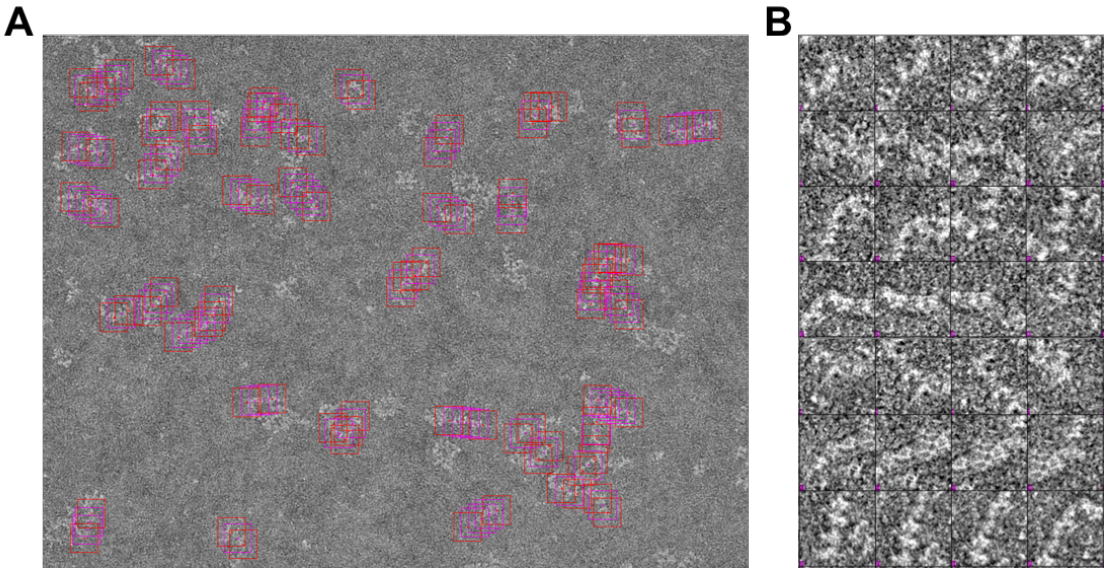


Figure 5.6. Particle picking from the heat polymer/4B12 Fab dataset. *An example negative stain micrograph of heat-induced AAT polymer with 4B12 Fab bound, imaged using a F20 microscope at 29,000X magnification. Movies comprising 18 subframes were collected over a 1.8 second exposure and summed. The summed image was CTF-corrected using CTFFIND3 and median filtered (shown) to aid particle picking. The polymers were boxed using EMAN1/Boxer with a box size of 200 pixels (pink squares) and a helix overlap of 140 pixels. B) Examples of boxed polymers.*

5.3.4.2 Iterative 2D alignment of picked particles

The picked particles were extracted from the micrographs for input into IMAGIC[121]. As at this stage they were all in different orientations, an initial alignment was required. The first alignment of the boxed particles against a low-pass filtered 55x200 pixel rectangle allowed rotational and translational shifts (Figure 5.7A). This rectangle size was chosen as several measurements of the diameter of the particles from the micrographs were 55 pixels (84.7Å). The alignment oriented all the polymers vertically through the centre of the box and required a large shift from the orientation they were picked in the micrographs. The particles were then subjected to multivariate statistical analysis (MSA) and classification, which groups and averages similar images. The class averages had an improved SNR compared to the individual images. They clearly showed that four subunits were present along the length of the box. In some of the class averages, the position of the 4B12 Fab could be discerned. Some of the best class averages that had good subunit delineation, but represented different views, were chosen as the references for multi-reference alignment (MRA) (Figure 5.7B).

The process of alignment, MSA and classification was repeated six times until satisfactory alignment of the tetramer had been achieved. Eigenimages which report on the major variations in the dataset were used to direct the alignments. For example, a horizontal striping pattern in an eigenimage indicates vertical misalignments between images; images were then classified based on such eigenvectors and aligned class sums or the total sum were used as references for a round of vertical image alignment. The amount of movement allowed in each iteration was gradually restricted to allow a more focused alignment. Horizontal alignment of the images was improved in a similar manner. Improvements in the alignments after several rounds were apparent from the total sums of the images, in which the subunits in the tetramer became better defined. Defined density for the Fab could not be seen in the total sums, but at this stage was not of concern as this would be averaged out from the different orientations. The total sum after round 6 was used to measure the size of subunits: the width of each subunit was on average 47 pixels, corresponding to 72Å, and the height was ~42 pixels corresponding to 64.7Å. The width of AAT is ~35Å in its shortest dimension and so the increased width in the class sum likely reflected horizontal misalignment and the presence of Fab, which has a length of ~60Å when viewed at an angle orthogonal to the variable loops. The

height of AAT is ~70Å which was in good agreement with the measured height from the total sum. Distances between subunits were also measured using the total sum by plotting grey values against distance in pixels: the average distance was 43.3 pixels, corresponding to 66.7Å (Figure 5.7C). This is in rough agreement with the expected distance of around 70Å found from undecorated polymers.

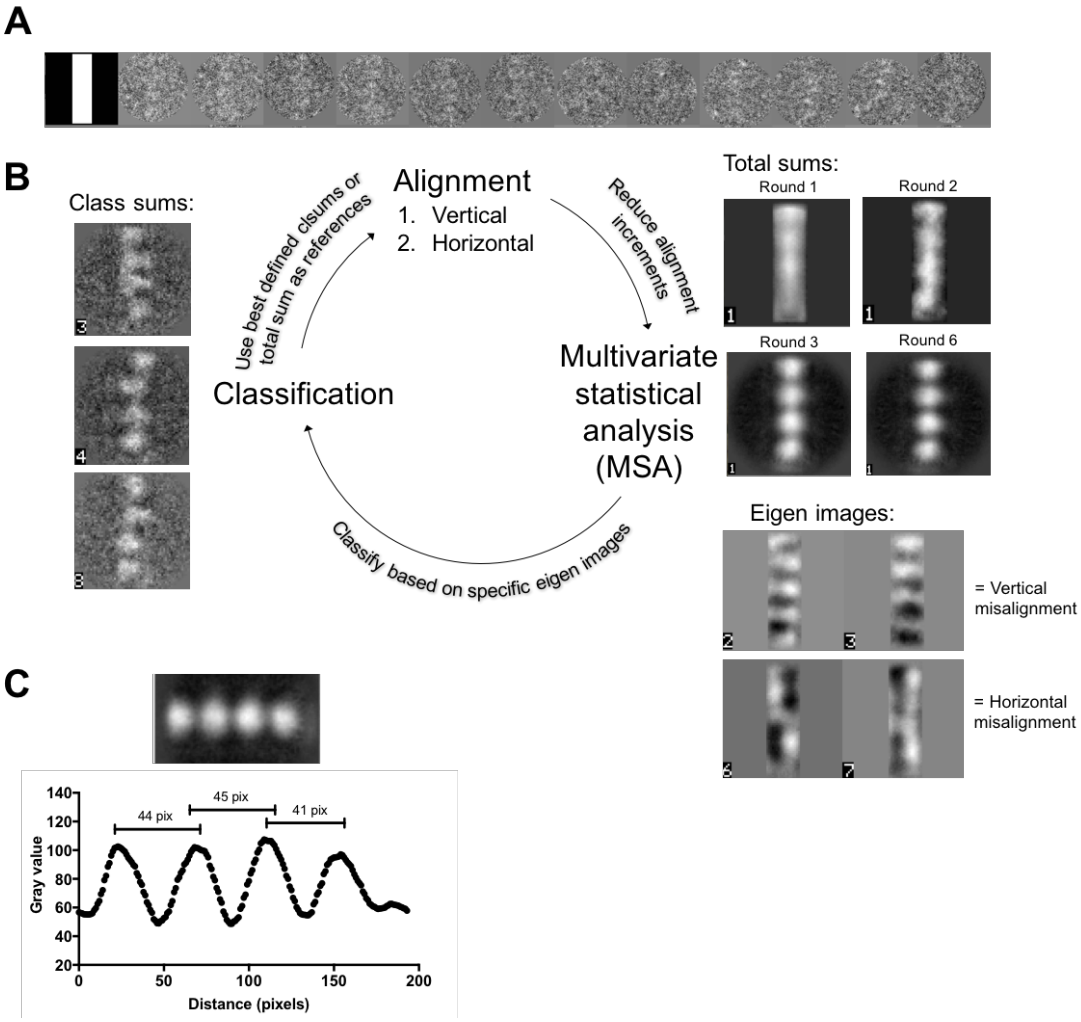


Figure 5.7. Initial alignment of the heat polymer/Fab particles. *Boxed particles were initially aligned to a vertical 55 by 200-pixel rectangle (top left) B) Particles were iteratively aligned: eigenimages produced from multivariate statistical analysis showed major variations in the data, these were used to generate class averages, which were then used as references in the next round of alignment. C) The class sum after 6 rounds of alignment was used to measure distances between subunits using grey scale analysis in FIJI*[257]*.*

5.3.4.2 Focusing 2D alignment on the AAT/Fab dimer

After horizontal and vertical alignment of the tetramer had been achieved, it was apparent from both the class averages and eigenimage analysis that the next major variations within the dataset was the presence or absence of the 4th subunit and the flexibility within the chain. Flexibility in particular would interfere with the optimum alignments of the images and consequently the accuracy of 3D reconstruction. Therefore, as before with the processing of the *ex vivo* NS-EM dataset, processing was focused on the smallest polymeric unit – the dimer – to limit the heterogeneity this produces. The aligned tetramers were trimmed to focus on the central dimer. The alignment-MSA-classification procedure was iterated again to improve vertical and horizontal alignments, as well as removing images that only had one subunit in the dimer. This left 12,823 particles.

Further heterogeneity was observed in the aligned dimer class averages. In some of the projection orientations with visible Fab electron density, it was clear that the AAT subunits either had a small angle between them (i.e. the Fabs were protruding from the same side of the AAT main chain) or a larger angle between them (for example there was Fab density on one subunit, but not on the other, or the Fabs were protruding from opposite sides of AAT). Combining these images for 3D reconstruction would also not yield a true structure. The images were classified according to eigenimages thought to represent this difference and then split by eye into two subsets. Examples of the class averages selected into the small-angle and big-angle subsets are shown in Figure 5.8. As there were side-view class averages where the AAT subunits had a near 180° rotation between them, images with two bright spots (only AAT visible, no Fab protrusions – marked by an * in Figure 5.8) were included in both subsets. Class averages that were difficult to assign to a subset were ignored. A total of 6,036 images belonged to the small-angle subset and 2,643 images belonged to the big-angle subset.

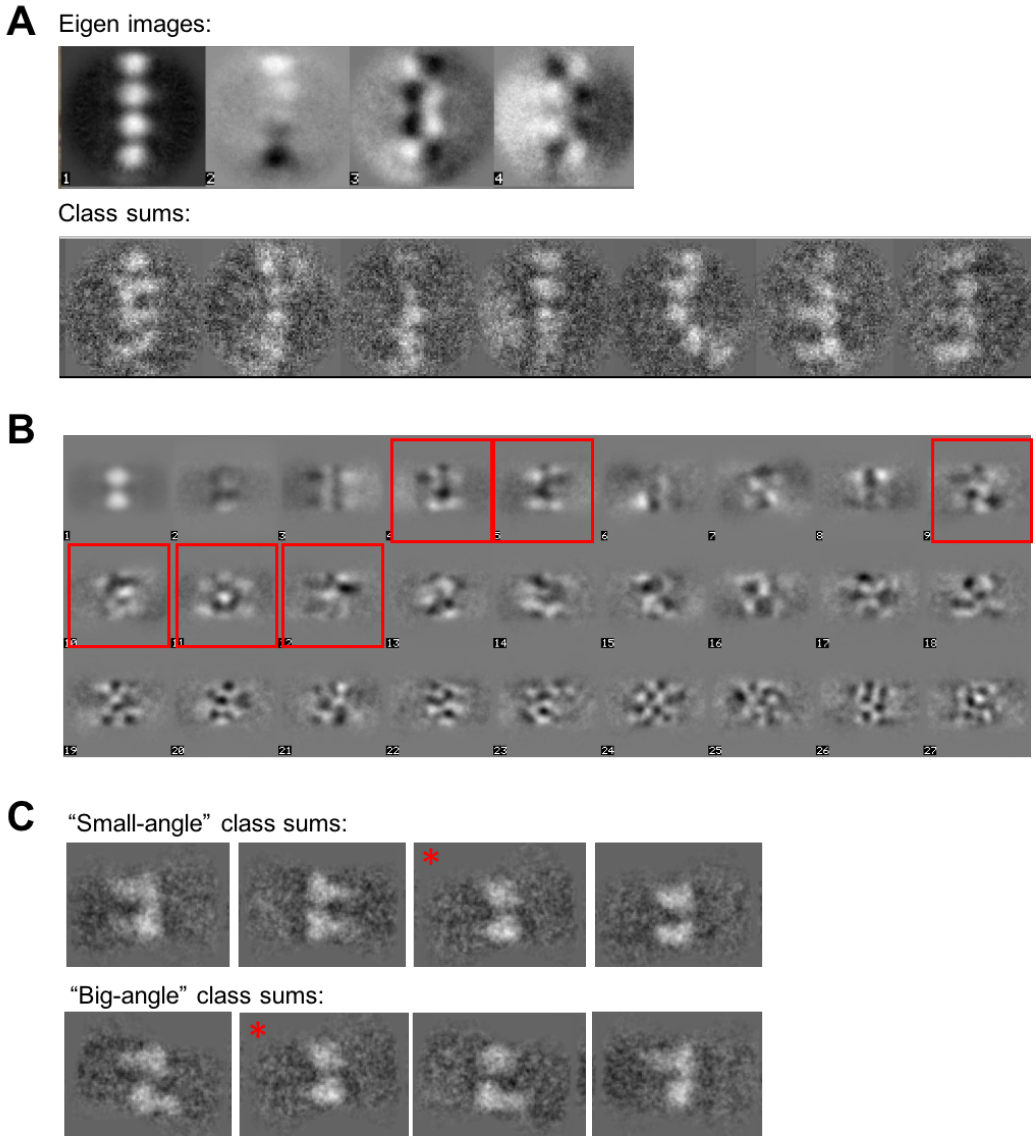


Figure 5.8. Eigenimage analysis of heat polymer/Fab particles. *A) Eigenimages and class averages of the tetramer showing the major variations are in the 4th subunit and in polymer flexibility. B) Eigenimages of the dimer where red boxes indicate variations in the Fab position. Images were classified into classes based on these eigenimages. C) Examples of the class sums that made up the small-angle and big-angle subsets. Class sums where the inter-subunit angle could not be determined (marked with an asterisk) were included in both subsets.*

5.3.4.3 Projection matching to generate a 3D structure

With a more homogenous set of images, it was possible to obtain a sensible 3D reconstruction of the small-angle subset. Angular assignment to 2D class averages was achieved using projection matching to references made from an initial model.

One of the *ex vivo* dimer reconstructions was used as the initial model. The elongated model with an 11° angle between the subunits was chosen as it had better-defined Fab density and had a linkage between subunits in comparison to the compact model (Figure 5.1). It was projected around the γ angle between 0-360° and around β between 70-110° (tilt) both in 10° increments to generate references that the images were aligned to (Figure 5.9A). This 'belt' of Euler angles was chosen because the polymers were flattened on the NS-EM grid and side-views of the polymer dominated the experimental projections. The generated 3D reconstruction of the best-matched class averages (as assessed by 3D-error) had well-defined density for both subunits and an angle of 17° between them. This model was then reprojected and matched to the images in an iterative process until convergence on 16° was achieved. The number of images within the class averages were gradually reduced from 25 to 10 and the references (reprojections from the model) were sampled at finer angles to refine angular assignment. The output models were assessed for angular distribution, visual comparison between reference reprojections and equivalent assigned class averages and model appearance (Figure 5.9C). Refinement was stopped after convergence had been achieved.

The final reconstruction was assessed by docking two copies of the AAT/Fab atomic model into the density using Chimera. The atomic model consisted of RCL-inserted AAT with the bound 4B12 Fab that had been modelled using the epitope mapping data as restraints[117]. Axes through the central masses of the AAT/Fab subunits were used to measure the distance and angle of rotation between subunits. The distance between subunits was measured to be 64Å and the rotation between the subunits was measured to be 16° (Figure 5.9B). The resolution of the 3D reconstruction was assessed to be 33Å or 26Å by Fourier shell correlation (FSC) with a cut-off of 0.5 or 0.143, respectively (Figure 5.9D).

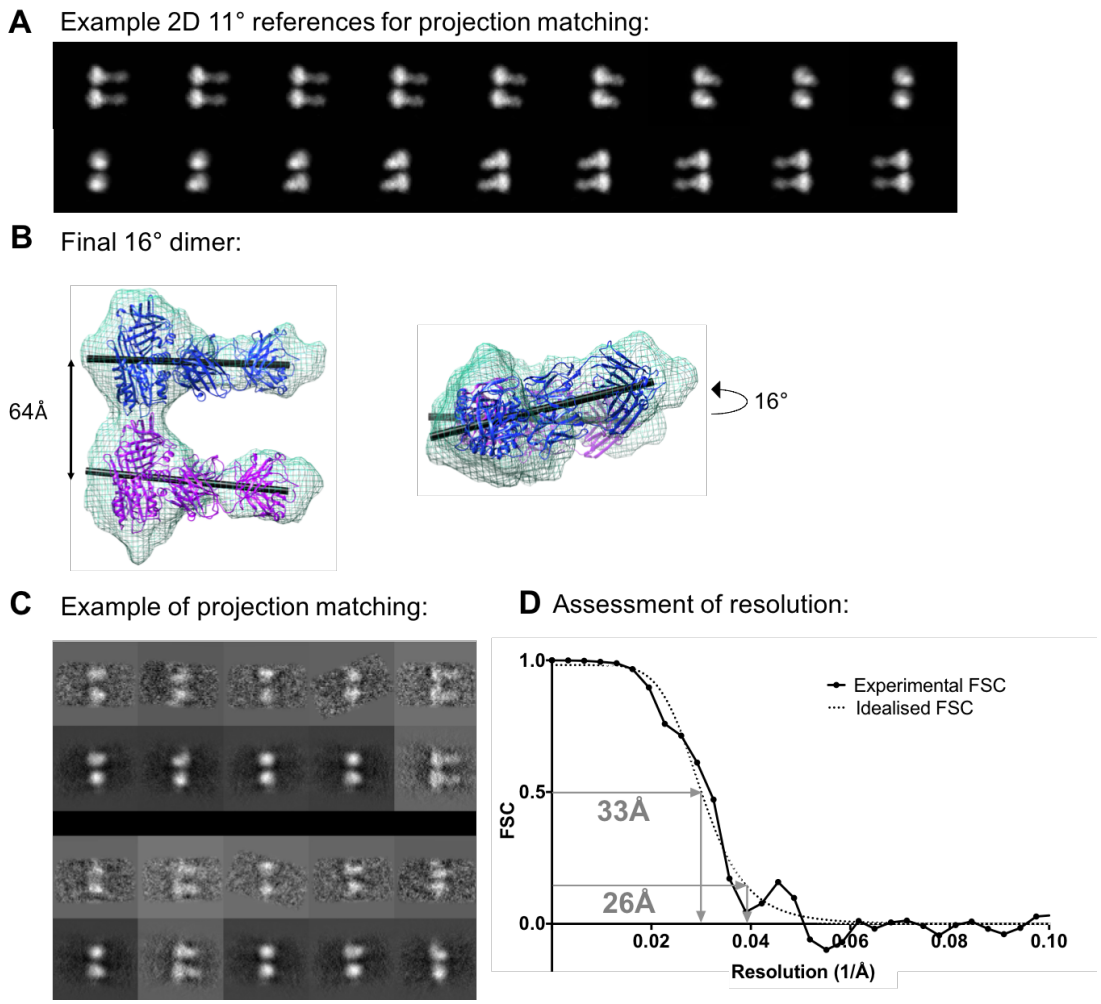


Figure 5.9. The 16° heat polymer/Fab dimer. *A) The 11° ex vivo dimer was used to generate references for projection matching against the experimental class averages. B) The final 16° dimer shown from side and top views. Two copies of AAT/Fab (in blue and purple) were fitted into the density using Chimera. Axes through their centres of mass were used to measure the inter-subunit distance of 64Å and rotation of 16°. C) Class averages (top) compared to their assigned reference projections (bottom). D) Resolution was assessed using Fourier Shell Correlation (FSC).*

To check for the possibility of model bias, the class averages with 10 particles per class were matched against the projections of multiple starting models of the dimer with different inter-subunit angles. Models were created in Chimera: the density of AAT/Fab subunits from the dimer model were superposed and averaged to create a single subunit model, the single subunit was duplicated and translated 70Å (based on an average measured distance from 8 class sums) and then one subunit was rotated in 10° increments. Models with a 0°, 10°, 20°, 30°, 40°, 50° and 60° subunit

rotation were made. Each of these models was then used as the basis for a single round of Euler assignment of class averages by MRA and subsequent 3D reconstruction. The angles between subunits in the resulting models, analysed in Chimera, differed by up to 15° from their starting point; the 50° MRA resulted in a model with a 35° subunit rotation and the MRA against the 10° and 20° models had the smallest angle difference. It appeared that there was a convergence towards an angle between 10 and 20°. This quick test suggests that the 16° angular information is representative of the data and the final angle is not dominated by bias from the reference.

The big-angle subset was also used in projection matching in an attempt to generate a 3D reconstruction. However, no suitable initial model existed for a dimer with a big angle between subunits. Initial models were created in the same way as above by rotating one subunit in 10° increments between 90° and 160°. This generated 8 initial models that were reprojected to make references for a competitive multi-reference alignment to the class averages. The greatest number of matches were to references with a 140-160° rotation and so projection matchings to a 140°, 150° and 160° model were carried out. The 3D reconstruction generated from the matches to the 140° reference had poor density for one subunit, indicating an incorrect angle between subunits. The 3D reconstructions based on the 150° and 160° references had a better second subunit and measured inter-subunit angles of 148° and 145°, respectively. Projection matching to a 145° model was carried out but the second subunit became less well-defined in the 3D reconstruction. It was therefore deemed not possible to get an accurate big-angle reconstruction, most likely due to the small number of images going into the reconstruction and the remaining heterogeneity between images.

5.3.5 Image processing of heat polymers bound with 4B12 by Relion

The same NS-EM dataset of heat polymer with bound 4B12 Fab was processed using the program Relion, to assess whether other 3D reconstructions were present within the data and/or validate the model obtained by IMAGIC. This analysis was carried out in parallel to the Relion processing of the *ex vivo* polymer with bound 4B12 Fab, which had also previously been analysed using IMAGIC (5.1.2). Therefore, in some of the steps described in this section, there was use of models generated from the *ex vivo* data.

Relion is an image processing system created in 2012 that is primarily based on maximum likelihood (ML) principles[122]. It can be operated via the graphical user interface, which lists the major image processing tasks in a standard workflow or via the command line. Many of the algorithms behind the programs are based on those found in other software such as xmipp[122]. The processing tasks are carried out in Fourier space. K-means clustering with ML methods underlie the classification steps[258]. Relion has addressed the problem of sample heterogeneity with a 3D classification pipeline, where multiple 3D structures can be reconstructed at once[259]. Relion was intended to increase the automation of image processing and be more accessible to the less-experienced electron microscopist[122]. This means some parameters are pre-set, minimising decisions of the user but with a concomitant loss in processing flexibility.

5.3.5.1 Autopicking and 2D classification of dimer images

The CTF-corrected micrographs used above were subjected to automated particle picking within Relion. References used for autopicking were generated from an early model in the *ex vivo* processing workflow (described below), made from a small sample of ~1000 particles and hence with little subunit definition (Figure 5.10A). It provided an adequate template to pick dimer particles. This model was projected around the Euler sphere to generate references of different views and maximise the chance of picking particles in all orientations (Figure 5.10A). The autopicking threshold used was low, again to pick as many particles as possible and the minimum inter-particle distance was set to 60Å to pick adjacent dimers from within the polymer chain (subunits are ~70Å apart). Picked particles were extracted in a 150-pixel box (230Å) and masked with a circle of 190Å diameter to allow space for Fab density either side of the polymer chain. They were then subjected to several rounds of 2D classification and selection to separate dimer particles from background noise (picked due to the low threshold) and single subunits. A set of 25,323 dimer particles was the result. The class averages had an average distance of 73Å between subunits and, based on the position of the Fab fragments, showed that the particles were of different orientations (Figure 5.10B).

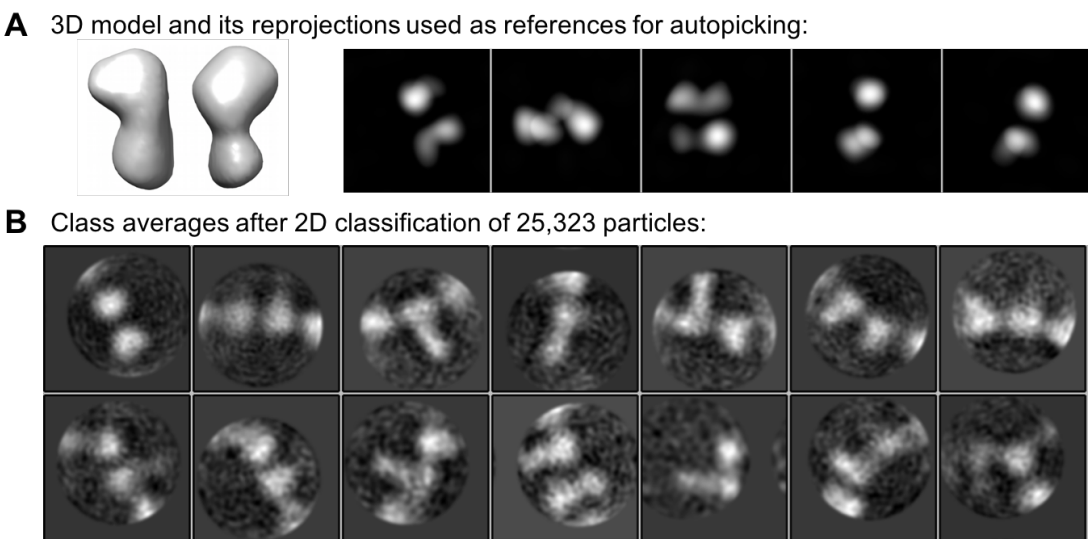


Figure 5.10. Autopicking and 2D Classification of heat polymer/4B12 Fab dimer particles using Relion. *A) The low-resolution model that was reprojected to produce example references for autopicking. B) Example class averages of dimers in a 150-pixel box and 190Å diameter circular mask.*

5.3.5.2 A reconstruction of the dimer using 3D classification to account for heterogeneity

From the 2D class averages generated in both IMAGIC and Relion, it was clear that the data were heterogeneous and there were differing rotations between subunits within the dimer. 3D classification attempts to address such structural heterogeneity by separating images according to the distinct 3D objects from which they arise[259]. The 25,323 particles were first sorted into 4 and then 8 models using a monomeric subunit (AAT/Fab) as a starting reference. A monomer was used as a starting reference to remove bias inherent in providing a dimer with a particular rotation between subunits. As such, if any of the resulting reconstructions contained a second subunit, this will have been built in purely based on the data. This reference came from one of the subunits of the Relion *ex vivo* polymer 3D dimer reconstructions and was filtered to 40Å. However, the 3D reconstructions generated were all poorly defined, with only one subunit formed, if any.

One of the Relion *ex vivo* dimer 3D reconstructions (discussed later in 5.3.6.3), which has a 60° angle between subunits, was then used as a reference in the generation of 5 models. Model 2 was the only model with two well-defined subunits (Figure 5.11A). However, as it was so similar to the input reference, it appeared likely that it was

formed as a result of initial model bias. A threshold was applied to this model, it was smoothened and low-pass filtered in Chimera and then used as a reference against the same 25,323 particles to produce several more 3D classes, again one of which was reasonable. This process was iterated to produce reconstructions more dependent on the data and less biased towards the first reference.

The 6,923 particles that comprised the best-formed class were refined against a reference of a single-subunit reference to ensure that the second subunit and the angle between them was inherent in the particles. A well-formed dimer was produced indicating that this was the case. The atomic model of the AAT-Fab complex was fitted into the two subunits and their central mass axes used to make measurements. There was a 67Å separation between the subunits and they were related by a 59° rotation (Figure 5.11A). The 3D reconstruction was masked with a soft-edge and the resolution assessed using FSC within the post-processing function in Relion. The resolution was 22Å or 27Å using a cut-off of 0.143 or 0.5, respectively (Figure 5.11B). An angular distribution plot of the particles constituting the reconstruction shows that side views around the 'belt' of the Euler sphere are well-populated whereas top views looking down the polymer axis are lacking (Figure 5.11C). To some extent, this was expected from the long length of AAT polymers and the flattening effect of negative stain. This anisotropic angular assignment however is likely to influence the reconstruction. For the purpose of verification, the 6,923 particles used in this final reconstruction were sorted by 2D classification to produce class averages (Figure 5.11A). From the position of the Fabs, it is clear the averages correspond to the refined model where the two subunits are related by a significant rotation.

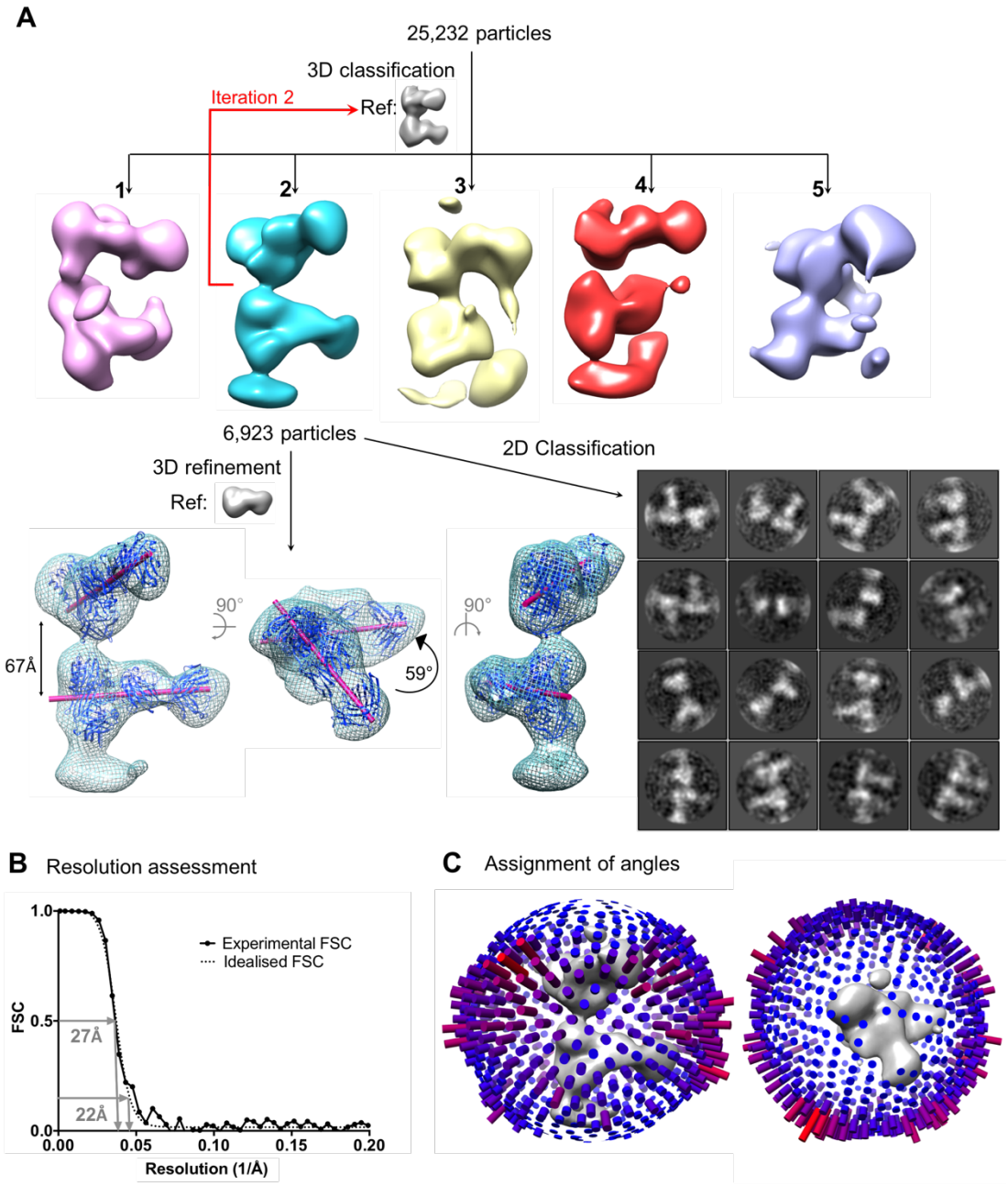


Figure 5.11. Workflow illustrating particle selection using 3D classification and the final 59° reconstruction. *A) 25,232 dimer particles were sorted into 5 reconstructions using the ex vivo 60° dimer reconstruction as a reference. The best reconstruction (Model 2) was then used as a reference in another round of 3D classification. The constituent 6,923 particles were refined against a monomer reference to produce the final map, displayed in side and top views. The molecular models of AAT in complex with 4B12 Fab (dark blue) were fitted into the structure using Chimera and the axis through the centre of mass (pink) used to make measurements. The 6,923 particles were visualised using a 2D classification. B) The resolution of the masked final model was assessed by FSC. C) A side and top view of the final model showing the angular distribution of projections that made up the reconstruction.*

To safeguard against model bias, the particles were used as input into an *ab initio* reconstruction in Relion. This function does not need a reference and so the produced reconstruction is purely from the data. Using all 6,923 particle images did not produce a reasonable reconstruction. This is not entirely surprising; while mis-assigned classes will be down-weighted during 3D classification and refinement, they can confound the *ab initio* algorithm. However, when the top 90%, 95% and even 99% of images (selected based on the probability score calculated for each during the 3D refinement) were used as input, after only five iterations, a similar model with a 68Å separation and 62° rotation resulted. This indicates the structure is real and highlights that a small proportion (1%, 70 particles) of junk particles can significantly influence the output in the *ab initio* reconstruction.

As a last verification, the 6,923 particles were classified by 3D into 2 or 4 classes against the filtered monomer reference. This was to assess whether any further separation of heterogeneity was possible. All of the generated reconstructions were reasonable in appearance. After fitting of the atomic models and making measurements, the distance between subunits only varied between 63-67Å, and the rotation varied between 56-60°. Therefore, it can be concluded that these particles cannot be divided any further and they are consistent with the 67Å separation and 59° rotation model generated.

Of interest, the 6,923 particles that contributed to the final reconstruction could not be refined against a reference of either one AAT subunit lacking a Fab or two AAT subunits without Fab domains. This suggests that without the Fab density as an anchor point, Relion struggles with the alignment of images. This is probably due to the featureless shape of the AAT molecule and emphasises the utility of the 4B12 Fab in providing asymmetry in the EM studies.

5.3.5.3 Attempts to identify more structures within the data

The particles that comprised the heat-induced 59° model were selected based on a 3D classification using the *ex vivo* 60° model as a reference. Therefore, the other models with different angular relationships were used as references in the 3D classification with the 25,232 particles, to select those belonging to other structures. Using the heat 16° IMAGIC model (5.1.2) or the *ex vivo* 88° Relion model (described in section 5.3.6.3) as references in 3D classifications, one out of the five output

structures bore resemblance to the reference (not shown). As before, the 3D classification was iterated, updating the reference each time with the best candidate from the previous round. However, when the images were selected based on the best model, they did not yield a dimer during refinement if provided with a monomer subunit reference, neither for the 16° nor 88° attempts (Figure 5.12). Both refined only to a sensible (but reasonably noisy) model with a dimer reference of their respective angles. However, this is likely due to a reference bias given a 2D classification of their constituent components was not consistent with the inter-subunit angles. The models were therefore discounted. The same process was carried out against the remaining particles not present in the heat 59° reconstruction (25,232 – 6,923) but the output did not change. The remaining particles were also used in a 3D classification against the monomer reference to try pull out other angular relationships but this also did not produce any reconstructions. Finally, a further selection of the best particles by 2D classification, followed by equivalent 3D classifications were attempted but were unsuccessful.

As a last attempt, the particles that were manually picked and comprised the heat 16° dimer IMAGIC reconstruction were entered into Relion. These 6,036 particles refined to some extent against a small angle dimer reference but never against the monomer reference, even after a further selection of particles (Figure 5.12). Therefore, the processing using Relion was not able to reproduce the IMAGIC 16° model in the absence of a dimer reference.

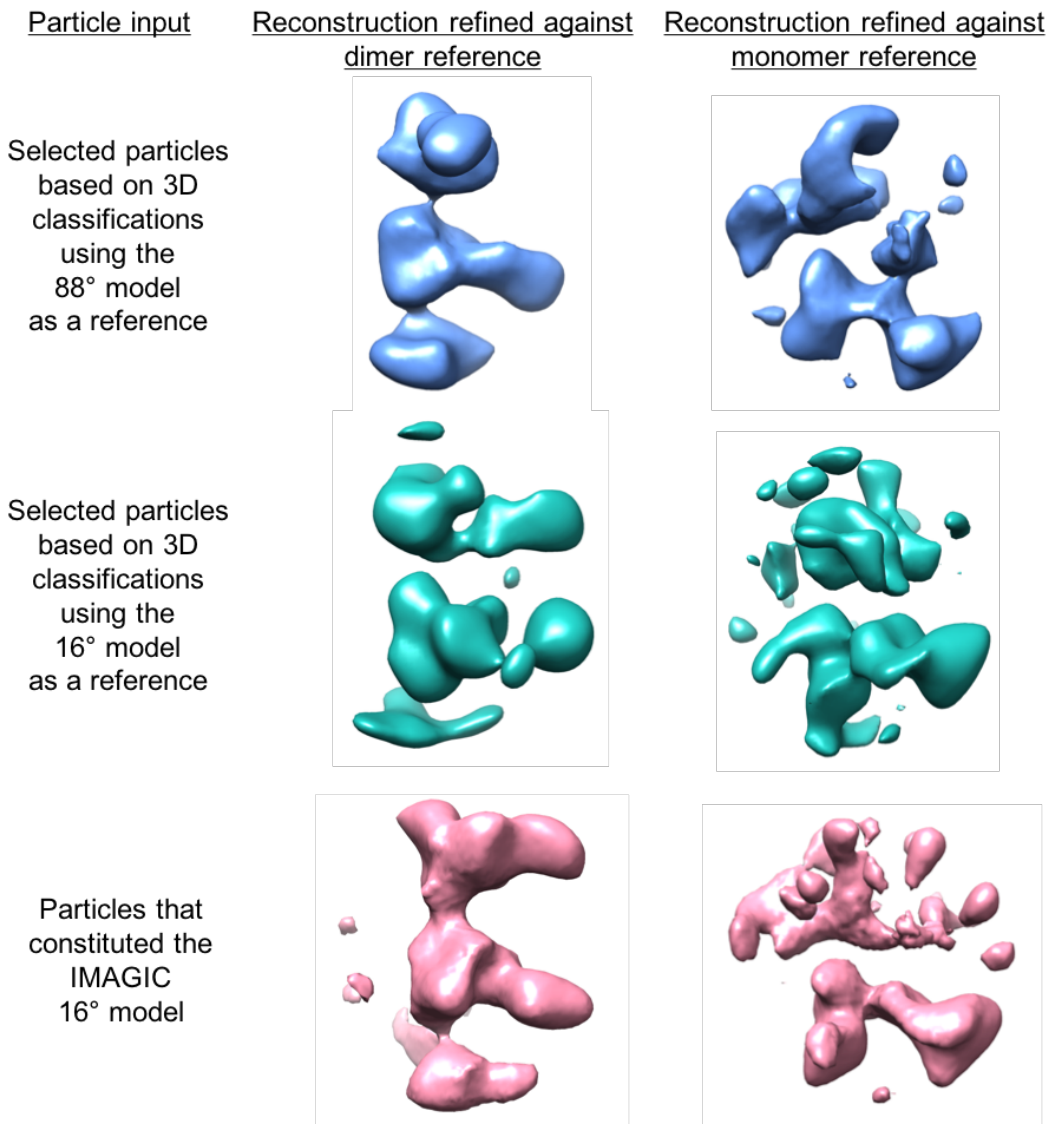


Figure 5.12. A comparison of structural refinements against a dimer reference or a monomer reference. *Attempts were made to find other reconstructions with differing angular relationships within the data. Images were selected based on a similar approach to that for the 60° model but the particles never refined to a sensible model with just a monomer reference. This was also true for the particles that made up the IMAGIC model.*

5.3.6 Re-processing of *ex vivo* polymers bound with 4B12 Fab using Relion

The NS-EM dataset of *ex vivo* polymers bound with 4B12 Fab that had yielded the elongated and compact 3D reconstructions of the dimer by IMAGIC processing, was also re-processed using Relion. Given the known heterogeneity of dimer

conformations, this was again to search for other dimer populations within the data. This was predominantly carried out prior to the re-processing of the heat dataset and therefore at points, a slightly different approach was taken.

5.3.6.1 Automated particle picking and classification of 2D particles

The Relion processing workflow was started at the automated particle picking stage as motion-correction and CTF-correction of micrographs had previously been carried out by MotionCorr2 and CTFFIND3, respectively[253]. The 11° *ex vivo* dimer generated by IMAGIC was reprojected at angles distributed around the Euler sphere and these projections were used as references for autopicking. The picking threshold was low, yet only a small number of particles (~3500) were selected. This was most likely due to the disparate grey intensity scales between IMAGIC and Relion; it is advised that reference-based auto-picking in Relion works best from references generated from the same dataset[139]. Nonetheless, the small number of particles were classified into 25 2D classes with a 180Å circular mask and the best classes representing dimers selected (~2000). These were used as input into a 3D classification job using the 11° model as a reference, where one of the models generated was in Figure 5.10A. It had little subunit definition but had two subunits an appropriate distance apart and was now on the correct intensity scale for another attempt at autopicking. Its projections (Figure 5.10A) were used to automatically pick particles from the whole dataset, again with a minimum inter-particle distance of 60Å to pick adjacent dimers and a low threshold to maximise particle numbers. It was reasoned that junk 'particles' could easily be identified and excluded by several rounds of 2D classification. A total of 176,957 particles were extracted into a 150-pixel box. Several rounds of 2D classification removed background noise and single subunit particles resulting in 68,538 good dimer particles (Figure 5.13). There were almost three times more dimer particles in this dataset than the heat polymer dataset, despite less micrographs, due to a higher concentration of polymer per micrograph. The class averages had two clearly defined subunits with an average distance of 74Å between them. A shape resembling the Fab could also be seen extending from the AAT chain on some of the subunits but interestingly, very few classes showed Fab density on both subunits in the dimer as would be expected from a side view of a small-angle dimer. To generate a 3D reconstruction that could be used as a reference for a 3D classification of the 68,538 particles, the very best dimer particles were further selected based on 2D classification, resulting in 15,911 particles.

5.3.6.2 3D classification rounds identify two different reconstructions

Several rounds of 3D classification were used to sort and separate the 2D particles into more homogenous subsets. First, a 3D classification job of the best 15,911 particles against the poorly-defined autopicking reference was carried out to generate four models. One of these (circled in Figure 5.13) stood out as having a well-formed bottom subunit with good shape complementarity to the AAT-Fab complex. The top subunit was stump-like, with no Fab density but the model was considerably less noisy than the others that were generated. This was the first model produced from the new particle set and it was hypothesised that the Fab definition of the top subunit may be lost due to an incorrect angle between subunits in the reference. Therefore, this new model served as a more accurate reference in the subsequent round of 3D classification against the full 68,538 particles. Furthermore, the lack of the top subunit Fab meant that little reference bias for inter-subunit rotations was introduced into the classification.

Due to a higher number of particle images and the expected heterogeneity within the AAT system, eight models were generated (Figure 5.13). These varied greatly in appearance; some (models 1, 4) had excellent single subunit contours, some had subunits well-positioned but less nicely defined (models 2, 5, 7, 8) and others were extremely noisy (models 3, 5). All but the latter, which did not constitute a 'sensible' model, were selected for their corresponding particles. This was to maximise particle numbers for later rounds of 3D classification. The same process was repeated once more with the stump-like model and then once the particles were deemed good enough, the reference for 3D classification was changed to be a well-defined single subunit. This single subunit was created in Chimera by erasing density for the second subunit from the stump-like model. The single subunit completely removed second subunit bias from the reference and meant the appearance of the second subunit in the classification output was entirely from the data. A dataset of 30,979 particles that resulted from this sequence of classifications was divided into four models. Two of these (models 2, 4) had excellent definition of both AAT/Fab subunits within the dimer and differed in the relative angle between subunits. The other two (models 1, 3) had only one well-defined subunit, and a stump for the other. Therefore, from this dataset, two reconstructions of Fab-bound AAT dimers were generated.

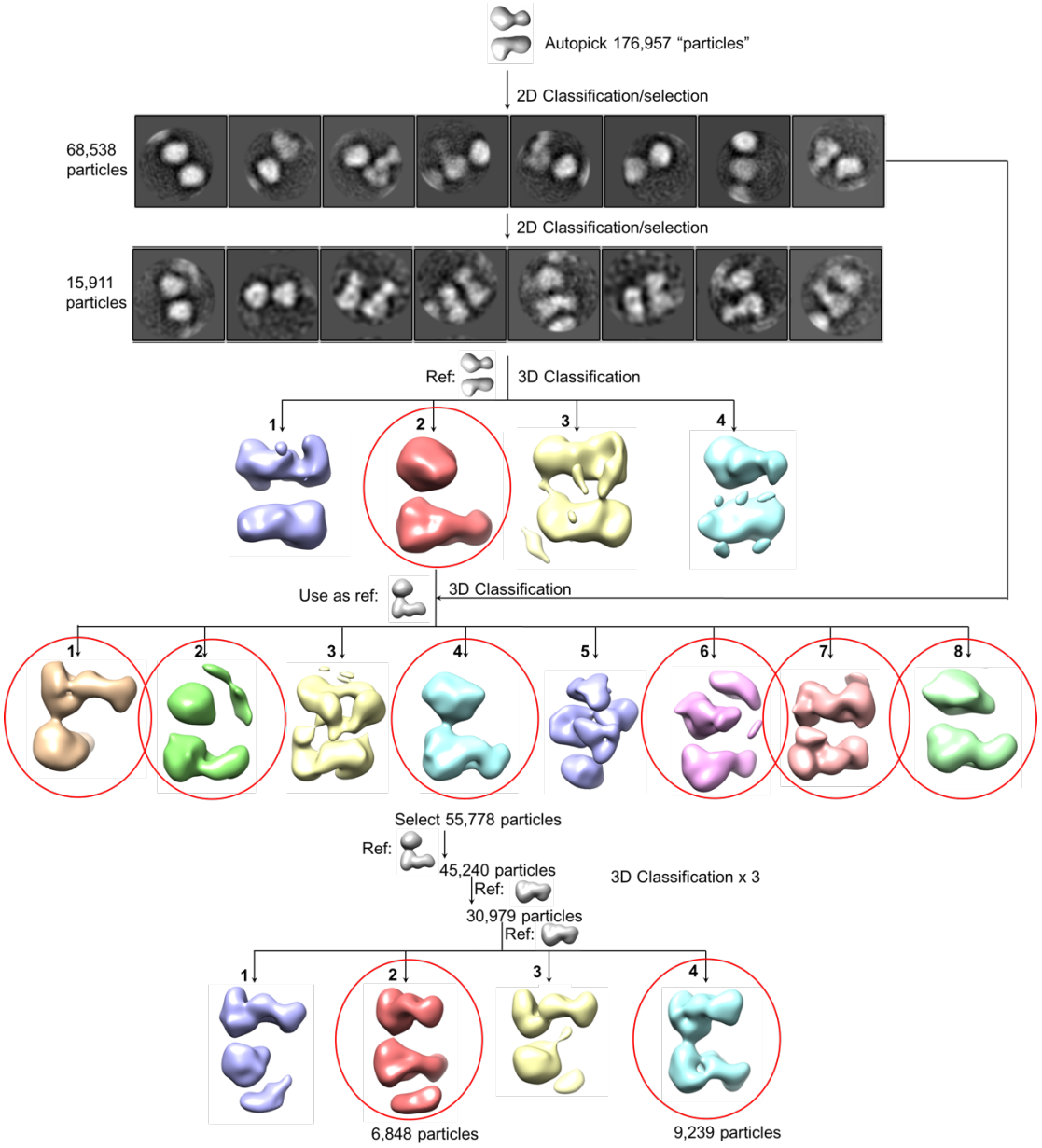


Figure 5.13. The complete processing workflow of the ex vivo polymer data using Relion. *Projections of a poorly-defined model were used to automatically pick 179,957 particles from the micrographs. Particles were ‘cleaned’ by several rounds of 2D classification and selection. 15,911 dimer particles were sorted into 4 reconstructions using the poorly-defined reference as a model. The best reconstruction (model 2) was then used as a reference for 3D classification against 68,538 particles. After several rounds of 3D classification, the strategy was tweaked, and a single subunit was used as the reference instead. From this, two reconstructions were obtained in which Fab density was clearly visible (model 2 and 4). All structures were visualised and checked for fits by a molecular model of the AAT-Fab complex using Chimera.*

5.3.6.3 Characterisation of the two reconstructions

The particles that corresponded to the two well-defined models were subjected to 'gold-standard' 3D refinement individually using the filtered single subunit reference, and the resulting final reconstructions analysed.

The first reconstruction resulted from the 6,848 particles belonging to model 2 (Figure 5.14A). As before, two copies of the atomic model of RCL-inserted AAT bound to 4B12 Fab were fitted into the dimer reconstruction using Chimera. Measurements were made using the axes through the centres of mass of the subunits. There was a 73Å distance between the two subunits in the dimer and they were related by an 88° angle. The post-processing function in Relion was used to add a soft-edged mask to the reconstruction and estimate its resolution. The FSC with a cut-off of 0.143 or 0.5 estimated a resolution of 22Å or 27Å, respectively. An angular distribution plot showed there was a strong preference for projections on one side-view face of the reconstruction. The 6,848 particles were also sorted by 2D classification to allow visualisation of the particles that constituted the reconstruction; it is clear the particles correspond to a dimer with a near 90° rotation between subunits.

The second reconstruction resulted from the 9,239 particles belonging to model 4 of the 3D classification step (Figure 5.14B). The subunit distance and relative angular rotation were measured to be 69Å and 60°, respectively. The resolution was estimated to be 16Å or 23Å using an FSC cut-off of 0.143 or 0.5, respectively. The increased resolution compared to the 88° model probably reflects the higher number of particles images making up this reconstruction and possibly particles with a higher informational content. An angular distribution plot showed the same preference for projections as for the 88° model but for this model the angles were determined more successfully and therefore were sampled on a finer grid. 2D classification of the constituent particles also showed from the positions of the Fabs that the angle was intermediate between planar subunits (0-20° rotation) and the near-90° rotation observed above.

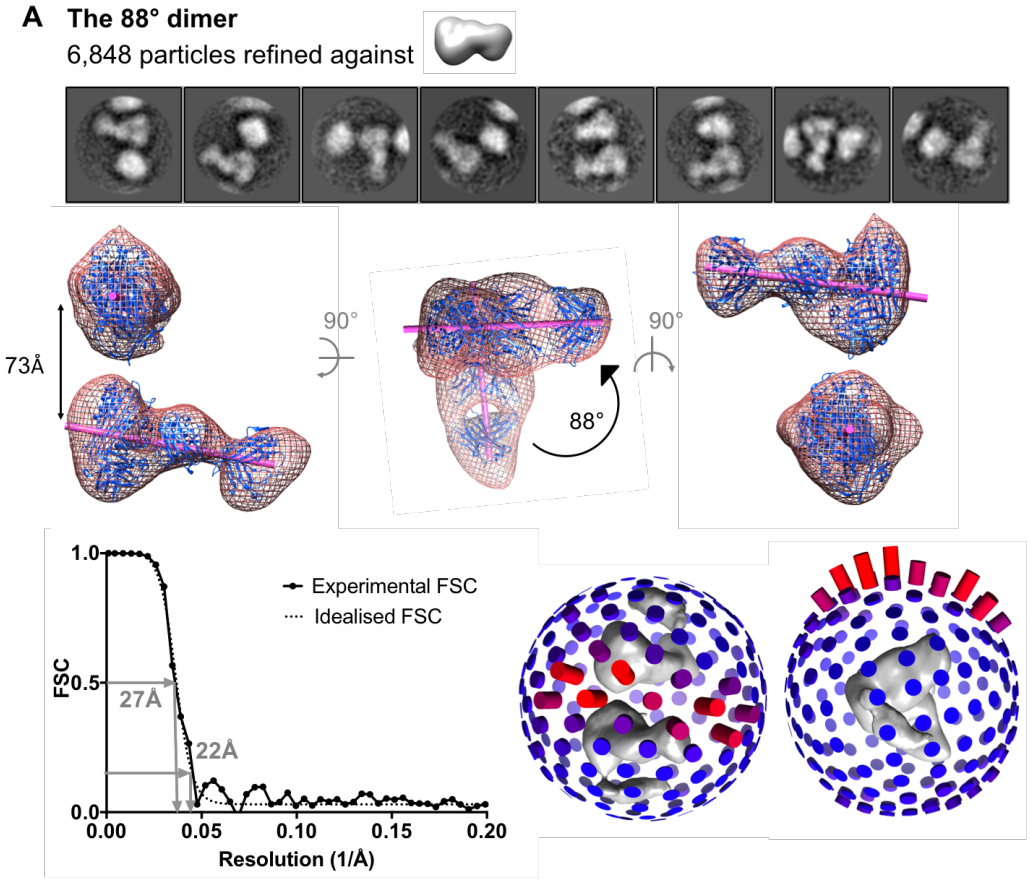

A
The 88° dimer
6,848 particles refined against
73Å
90°
90°
88°
1.0
0.5
0.0
FSC
27Å
22Å
Experimental FSC
Idealised FSC
0.00
0.05
0.10
0.15
0.20
Resolution (1/Å)
B
The 60° dimer
9,239 particles refined against
69Å
90°
90°
60°
1.0
0.5
0.0
FSC
23Å
16Å
Experimental FSC
Idealised FSC
0.00
0.05
0.10
0.15
0.20
Resolution (1/Å)

Figure 5.14. Characterisation of the two ex vivo dimer reconstructions. *The particles belonging to the best reconstructions after 3D classification were refined against a monomer reference to produce the final reconstruction, displayed in side and top views for A) the 88° dimer and B) the 60° dimer. The models of AAT/Fab (dark blue) were fitted into the structures using Chimera and the axes through the centres of mass (pink) were used to make distance and angle measurements. The constituent particles were also visualised by 2D classification. The resolution of each reconstruction was assessed by FSC and the angular distribution of projections is shown plotted on side and top views.*

Both reconstructions could almost be completely regenerated using five iterations of an *ab initio* reconstruction in Relion using their respective particle sets. For the 88° particles, the output of the initial model job yielded a reconstruction with a 73Å subunit distance and 78° subunit rotation. For the 60° particles, the output of the initial model job yielded a reconstruction with a 68Å subunit distance and 58° subunit rotation. This gives further confidence that the 2D particles are true projections of their 3D reconstructions. In comparison to the Relion processing of the heat polymer, no filtering of particles based on the probability score was required. This might indicate the particles were of a higher quality or that less 'junk' particles were present in the liver polymer dataset influencing the refinement.

The particles constituting each reconstruction were also subclassified by 3D into 4 classes against the single subunit reference to assess homogeneity of the particles. For the 88° particles, the reconstructions were similar but with a degraded quality. The distances between subunits varied between 66-74Å and the rotation angle varied between 75-86°. The same was true for the 60° particles with the distances varying between 67-70Å and the rotation angle varying between 59-68°. These results concluded that there was no further subclassification of particles and the particles are consistent with their respective models. It is expected that these inter-subunit angles are enriched components of a continuum, which would suggest that some of the constituent particles arise from dimers with rotations distributed around these central points.

5.3.6.4 Attempts to obtain other reconstructions within the dataset

The 14,892 particles (30,979 – 6,848 – 9,239) that did not make up the two final reconstructions were further tested to see whether another reconstruction was within

the data. They were reclassified against the monomeric subunit reference into between 2, 4 and 10 classes (Figure 5.15A). In all the generated classes, a second Fab-bound subunit was not obtained at a different angle to the already determined reconstructions; they all had a well-formed single subunit and a stump for the second. Therefore, no further models with different angular relationships could be found. The 60° and 88° model did appear again, indicating that firstly, not all the particles corresponding to their projections had been identified and secondly, that the previously-determined reconstructions were credible. Projections of the 60-88° dimers are apparent in the class averages of the remaining particles (Figure 5.15A).

The 14,892 unassigned particles were also classified in 3D against the previously-obtained IMAGIC-derived reference map with a small angle (11°) between subunits (not shown). Some of the generated models had a small angle between subunits but when images were selected from the models on this basis, they did not refine against a monomeric subunit. This suggests the appearance of a small angle dimer is largely due to reference bias. Even when the initial 68,538 particles were classified against the small angle dimer reference and the same selection process repeated, no unbiased small-angle model resulted.

The 8,120 particles that were manually picked and formed the two small angle models using IMAGIC were also tested in Relion. They were refined against the monomeric subunit but a sensible model was not produced (Figure 5.15B). They even did not yield a good model when refined against the 11° IMAGIC model. Further selection of images and updating of the reference from job outputs did not make a difference. When the particles were analysed by 2D classification, they were well-defined but some resembled projections of a model with an angle more similar to 60° or 90°, rather than 11-16° (Figure 5.15B).

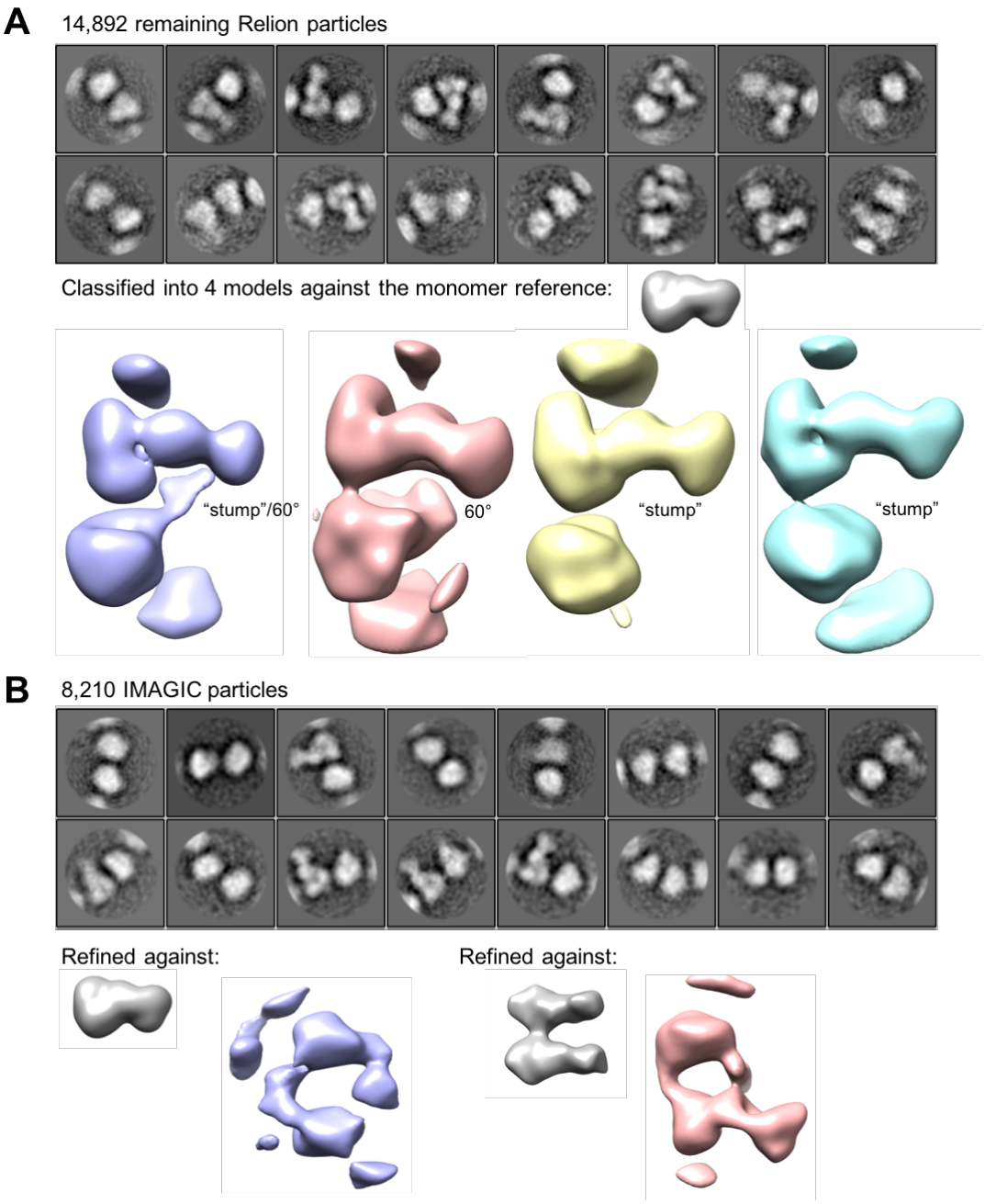


Figure 5.15. Attempts to identify more ex vivo reconstructions. *A) Unassigned Relion particles were visualised by 2D classification and subjected to 3D classifications. No new angular relationships were identified in the generated models. B) The manually-picked IMAGIC particles were visualised by 2D classification and failed to refine to a sensible model.*

5.3.7 Refining the subunit model of AAT/Fab using the single subunit density

The NS-EM reconstructions generated by Relion had greater subunit definition, which allowed the model of the interaction between 4B12 Fab and AAT to be

improved. An average subunit density was made from the two subunits in the Relion 60° dimer reconstruction as this was the structure with the highest resolution. The interaction between cleaved AAT (PDB ID:1EZX[13]) and the 4B12 Fab was refined using PyRosetta with the help of Dr J. Irving[164]. A number of "rule-based" restraints were included in the modelling: residues 32, 306, 43, 36, 266 in AAT (identified from epitope mapping studies) had to be within interaction distance of the 4B12 Fab and the heavy *or* light chain CDRs of the Fab had to be within interaction distance of AAT. A set of 1000 randomised starting models of AAT and Fab were optimised and fit as rigid bodies into the density. The resulting subunits models were sorted based on the cross-correlation score measuring their fit to the NS-EM single subunit density. A histogram of the cross-correlation scores shows the distribution of models was bi-modal (Figure 5.16A). Models with lower cross-correlation scores were visibly worse in their fit to the electron density; examples are shown where the Fab region is not in the density or the structure is flipped around such that AAT is in the density for the Fab and vice versa. In the top 20 structures (the highest scoring bin), with the exclusion of one model where the AAT molecule was wrongly positioned in the Fab density, all models favoured the Fab heavy chain and light chain in particular orientations (Figure 5.16B). This suggests there is structural discrimination in the Fab region of the electron density. The orientation of the AAT molecule was not so easily discriminated (13 'up' vs. 6 'down' orientations). Therefore, the highest-scoring of each AAT orientation was chosen – model 329 and model 787 (Figure 5.16C) – and used to create two different versions of each polymerisation model for the fitting studies described in the next section.

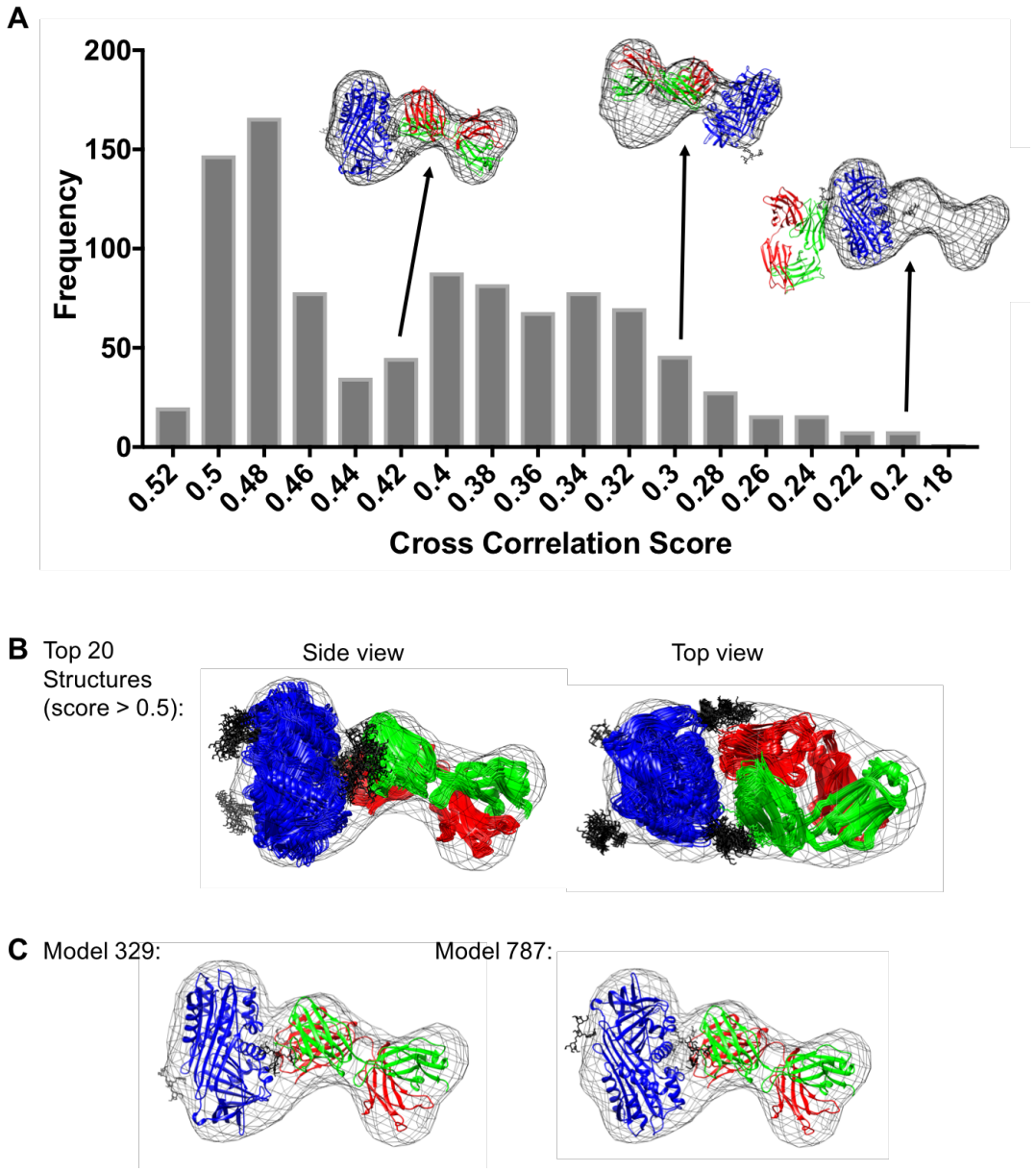


Figure 5.16. Improving the modelling of the 4B12 Fab/AAT interaction. *A) 1000 random models were created and fit as a rigid body into an average density of the single subunit. The cross-correlation score measured the fit to the electron density and the histogram shows the spread of model fits. Example models with differing scores are shown. B) The top 20 models with AAT in blue, the heavy chain in red and the light chain in green. Glycans are in black and the average density is thresholded at a volume corresponding to 1.5x the expected molecular weight. C) The best models with AAT in opposite orientations.*

5.3.8 Fitting the proposed polymerisation models into the NS-EM reconstructions

The loop-β sheet A, β-hairpin and C-terminal models of polymerisation have all been proposed to represent the linkage that exists between subunits in *in vivo* polymers[32,78,86]. The models are based on different biochemical and structural data of serpin polymers formed artificially but it is unclear if they represent the structure of liver polymers. The three polymerisation models were fit into the NS-EM reconstructions generated from the *ex vivo* polymer data, as well as those generated from the heat-induced polymer data to see whether they were compatible with the structures. It was the assessment of whether the polymerisation models could accommodate such an angle between subunits that was important, rather than the presence of such a dimer in the sample.

The three polymer models had been generated for fitting into the liver polymer IMAGIC models, previously[250]. Briefly, the loop-β sheet A model was modelled with an RCL insertion of 6 residues (P8-P3). This minimal linkage was chosen to maximise the flexibility of the loop-β sheet A model and likelihood of fitting the structures, but represents a departure from the fully-inserted RCL proposed in early papers which would result in rigid, truncated polymers[91]. The β-hairpin model was generated in a linear conformation, and in contrast to the published linear model, was produced with a folded I-helix based on the finding that this element is involved in 4B12 Fab binding[69,78,117]. The C-terminal model was also generated in a linear conformation by altering the connectivity of the structure of the cleaved polymer of AAT[86]. The AAT/Fab subunit models generated in 5.3.7 (model 329 and model 787) were used to redefine the AAT/Fab interaction in the polymerisation models. All the NS-EM density reconstructions (60° liver, 88° liver, 16° heat, 59° heat) were compared against both a 329-version and a 787-version of each polymerisation model. Both subunit models were used because the direction of AAT in the density could not be accurately determined in the fitting of the single subunit.

A molecular modelling protocol was developed to perform rigid-body modelling of the AAT/Fab subunits, with defined flexibility in the connecting polymerisation linkers. This was carried out in PyRosetta[164] with the help of Dr J. Irving. A total of 500 randomly perturbed starting models were generated, optimised into the densities and given a score based on the cross-correlation of the model to the NS-EM structure.

The starting models covered a wide range of conformations between subunits and the highest 10 scoring structures were all similar with a small standard deviation between scores. This showed that in each discrete run, the models had converged onto a similar solution, which gave confidence in the robustness of the modelling. The average cross-correlation scores of the top 10 models for each polymerisation run are shown in Table 5.2.

		Version	Loop-β sheet A model	β-hairpin model	C-terminal model
Liver polymer	**60°**	329	40.55	44.08	46.71
		787	33.88	43.39	45.25
	88°	329	43.34	43.43	46.41
		787	35.80	42.25	45.05
Heat polymer	**16°**	329	32.25	31.70	34.01
		787	27.11	32.51	32.16
	59°	329	42.41	43.24	46.58
		787	30.50	42.33	43.74

Table 5.2. Correlations of the fits of the polymerisation models into the NS-EM densities. *Modelling was carried out using PyRosetta and correlations were calculated between the models and the NS-EM structures. Values are the average of 10 models, all runs had a standard deviation of ≤ 0.03. The highest score for each structure are highlighted in grey.*

The correlation scores are a measure of the ability of the different models to fit the density. They reflect the restraints conferred by the different linkages on the translational and rotational flexibility of the individual subunits. For all four NS-EM structures, fit with a specified version of the polymerisation model (i.e. 329 or 787), the C-terminal model of polymerisation always had the highest score, with one exception, and hence best fit the density. The loop-β sheet A model always had the lowest score and the β-hairpin model, the intermediate score. The fitting was also able to suggest the most compatible orientation of AAT in the densities. When comparing the scores between the 329- and 787-versions of the polymerisation models for each structure, the 329-version always had the highest score. Especially

for the loop-β sheet A model, the difference between scores was substantial. This suggests that the direction of AAT can be determined when modelling into the dimer structures due to the extra constraints of the neighbouring subunit (in contrast to the modelling into the single subunit structure in 5.3.7).

An evaluation of whether the fits of the different models were statistically different was undertaken. The original 329 or 787 individual subunits were superimposed over the two subunits in each of the 10 top-scoring dimers for each polymerisation model to calculate a starting cross correlation (CC). The superimposed subunits were then optimised into the respective density *without* any restraints imposed by the polymer models to calculate a resulting unrestrained CC. This approach was taken as the polymerisation models involve different structural elements (particularly at the termini) which can lead to intrinsic differences not directly related to the fit. The restrained CC (from the optimised dimer models) was calculated as a fraction of the unrestrained CC (from fitting the individual subunits) for all the model combinations. Two separate non-parametric tests of significance were carried out, testing whether the individual subunits were significantly different to the best-fitting polymerisation model. This effectively tested the ability of a polymerisation model to converge on an optimal solution despite its restraints. The first test concerned only liver polymers and assumed a common mechanism resulted in both the liver 60° and the liver 88° structures. The 60° and 88° fits were combined into one score for 329 or 787, and the highest similarity dimer combination was the C-terminal (CT) 329-model. Applying the Dunn's non-parametric multiple comparisons test, there was a significant difference ($p < 0.0001$) between the CT-329 model and all the β-hairpin (BH) and loop-β sheet A (LS) models (Table 5.3, left). The second test assumed that a common mechanism resulted in the liver and the heat structures and the scores were combined accordingly. The highest scoring dimer combination was again the CT-329 model and this was statistically different ($p < 0.0001$) to all the BH and LS models (Table 5.3, right). Furthermore, CT-329 was also significantly different ($p < 0.01$) to the CT-787 model. Overall, this analysis suggests that the C-terminal model of polymerisation fits significantly better into the densities than the other models of polymerisation.

Liver dimer comparison			
	LS	BH	CT
329	**** 0.891	**** 0.933	0.997
787	**** 0.769	**** 0.938	0.995

Liver/Heat dimer comparison			
	LS	BH	CT
329	**** 0.891	**** 0.927	0.998
787	**** 0.728	**** 0.939	** 0.985

Table 5.3. Dunn's non-parametric multiple comparisons test evaluating the polymerisation model fits into the NS-EM densities. *The 60° and 88° liver dimers were assumed to result from the same polymerisation linkage on the left, and the liver dimers and heat dimers were assumed to result from the same polymerisation linkage on the right. The similarity values represent the restrained CC value as a fraction of the unrestrained CC value. LS = loop-β sheet A model, BH = β-hairpin model, CT = C-terminal model, **** denotes significance p < 0.0001 and ** denotes p < 0.01.*

The order of scores was also apparent when visually assessing the fits; the fits into the liver 60° model as a representative is shown in Figure 5.17. For example, the linker in the loop-β sheet A model is constrained such that one AAT subunit (blue) is pulled out of the main body of electron density. The linker is already minimally-inserted (only 6 residues) and therefore the loop-β sheet A model as originally proposed with a complete RCL swap, would be incompatible with all structures. The 6-residue swapped version can broadly fit the structures but considering it is the lowest scoring model, and the minimal-insertion is unlikely to confer the high known stability of AAT polymers, the findings suggest this linkage may not be correct. The modified β-hairpin model cannot be discounted based on these fits as it could reasonably accommodate the geometric restraints provided by the structures and resulted in an adequate score. Although the C-terminal model is the most flexible model and therefore it is inherently most likely to fit the structures, these findings suggest that the reconstructions of the liver polymer and the heat polymer are best described by the C-terminal linkage of polymerisation[86]. A definitive analysis however would require higher resolution structures.

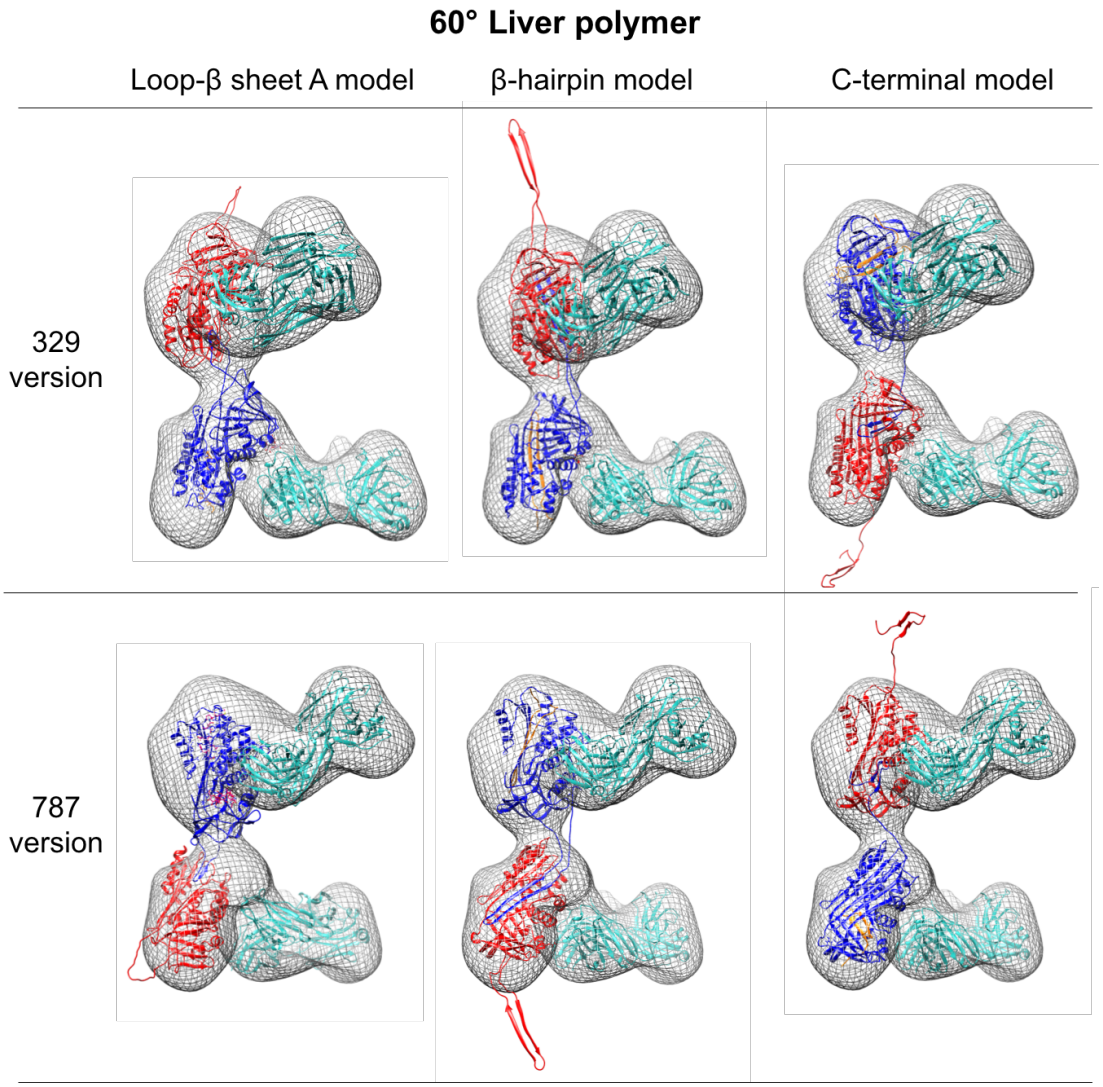


Figure 5.17. Fits of the NS-influenced polymerisation models into the 60° liver polymer structure. *AAT subunits are shown in red, blue with the neighbouring AAT linkage in orange and the Fab subunits are in teal. Modelling was carried out using PyRosetta and visualised using Chimera.*

5.4 Discussion

5.4.1 Comparing the structure of heat-induced polymers and *ex vivo* polymers

5.4.1.1 The two polymer types have similar ultrastructure

Heat-induced *in vitro* polymers and liver/*ex vivo* polymers, but not *in vitro* polymers formed by other methods (acid and denaturants), are recognised by the mAb 2C1[76]. This indicates that heat polymers and liver polymers share a common epitope and therefore share at least some degree of structural similarity. However, the extent of this similarity has not been investigated.

Negative stain EM (NS-EM) is a convenient means by which to visualise protein samples and has been used previously to visualise AAT polymers. Polymers isolated from the liver of a patient undergoing transplantation for Z-AAT cirrhosis, plasma-purified Siiyama AAT and heat-induced polymers have all been reported to consist of flexible chains with bead-like subunits[32,81]. This morphology has been observed also in this study when considering heat-induced polymers (Figure 5.3) and *ex vivo* polymers isolated from the liver of a Z-AAT homozygote (Figure 6.6), with a view to extending these analyses to infer structural information. In the images that have been collected, individual subunits can be clearly defined, with polymer chains ranging from 4-20 subunits. For the heat polymer sample, linear and circular polymers can be seen, both present in significant amounts, whereas fewer circles were observed in the *ex vivo* micrographs (Figures 5.3 and 6.6). In fact, a quick assessment that classified over 2000 polymers, from both the heat-induced and liver samples, as circular or linear showed that there was an 81%/19% split of linear to circular polymers in the liver polymer sample compared to 67%/33% in the heat polymer sample. Polymers are known to have complementary ends as circular polymers have been reported in the *ex vivo* samples before, however only as the minority species[32]. While it is possible that these circular values are over-estimations – there may be linear polymers that have adopted a circular shape – the number of circular polymers counted here, particularly for the heat polymers, represents a substantial component. As polymerisation involves inter-molecular interactions, it is by definition a kinetic process, and there would be competition between monomer addition and circularisation. It is likely that the means by which polymerisation is induced, the

presence of mutations, and the concentration of the components would all influence how well-favoured elongation is over circularisation. Thus, studies that draw conclusions regarding the importance of circularisation must be interpreted with care[86,95].

The featureless shape and small size of the AAT subunit makes alignment of 2D projections and generation of a 3D reconstruction of the polymer challenging. Binding of the 4B12 Fab to the polymer not only increases the size of the repeating unit but as the antibody binds a specific epitope, its position provides information on subunit rotation relative to each other. Binding of the 4B12 Fab to liver *ex vivo* polymer facilitated the generation of 3D reconstructions of the dimer in previous work[250] (Figure 5.1). The 4B12 Fab was bound to heat polymers for visualisation by NS-EM (Figure 5.5), and to the *ex vivo* polymers for cryo-EM work (Figure 6.7). The bound 4B12 Fab molecules increased the size of the polymers and were often visible as protrusions from the polymer chain. Whether complete subunit binding was achieved is difficult to assess from the micrographs, as the 2D projection shape of orientations orthogonal to the plane of the micrograph would appear the same if Fab was or was not bound. The intensity of such subunits does often seem to be brighter however, suggesting an increase in the protein component. The sample was incubated with three times excess of Fab for multiple hours, 4B12 has a sub-μM affinity for AAT (preliminary data from A. Jagger), and visual inspection of the chains provides confidence that occupancy is at least very high.

The flexibility of the heat polymers was immediately apparent from the micrographs and further exemplified by eigenimages and class sums of tetramers (Figure 5.8). This flexibility and heterogeneity between images led to the choice to limit reconstruction to the dimer. In general, the flexibility of the polymer chain is comparable between the two polymer types. In some regions, the protruding Fab was lined up on one side of the chain whereas in some instances the Fab could be seen alternating on opposite sides of the chain. This latter case indicates the substantial flexibility of the inter-subunit linker, in which one monomer rotates approximately 180° relative to its neighbour. It is clear however that this latter arrangement is observed substantially less frequently, suggesting an angular preference. Previous work using NS-EM to study the *ex vivo* polymers with 4B12 Fab found the Fab protrusions nearly always on the same side of the polymer chain[250]. When a different liver sample was characterised by NS-EM in preparation for cryo-

EM, the majority appeared to be on the same side but a few cases with alternating sides could be observed Figure 6.7). The difference between the two liver polymers may indicate heterogeneity between liver polymers from different patients[102], or simply that they are permissible, but generally disfavoured relative rotation between subunits. Regardless, these extreme angles of rotations are apparent in both heat-induced and liver polymer samples further supporting their similarity at an ultrastructure level. It is also important to remember that staining with uranyl acetate has a flattening effect[129] and so if subunits are rotated by greater than 90°, it is possible that they could be flattened and appear like a 180° rotation.

5.4.1.2 The two polymer types display similar angular relationships between subunits

In this work, two NS-EM reconstructions have been determined for the 4B12-bound heat polymer sample using two different software approaches and two new reconstructions in a re-appraisal of the liver polymer sample. In combination with previous work primarily carried out by S. Faull[250], these studies present the first EM structures of AAT polymers and the most thorough comparison of heat-induced and liver-derived polymers. The reconstructions are summarised in Table 5.3.

A dimer structure of the labelled heat polymer was generated using IMAGIC with a 64Å distance and 16° rotation between subunits. This was very similar to the IMAGIC models from the liver polymer dataset generated previously with respective parameters of 68Å/16° and 73Å/11°. The other heat dimer determined using Relion had a 67Å distance and a 59° rotation between subunits. At the resolution of the structures, this is indistinguishable from the liver dimer newly-determined using Relion with a 69Å distance and 60° rotation. Therefore, comparable reconstructions of dimers are observed between the two polymer species. As discrete angles are observed, as opposed to a continuum of angles, this suggests the inter-subunit linkage and/or the negative staining process favours certain subunit positions. This angular preference is common between the two polymer species and supports studies claiming their structural similarity[69,76].

The only different dimer species that was identified, was another liver-derived dimer with a 73Å distance and 88° rotation; no equivalent structure was found in the heat dataset. As the heat-induced polymer dataset contained around three times fewer

particles than the liver polymer dataset, this angular relationship may simply not be present in high enough frequency to form a reconstruction. An alternative is that the near-90° angle is not a favoured or permissible rotation between subunits in heat-derived polymers, possibly reflecting a difference in the polymer linkage between heat-induced and those derived from liver. This seems an unlikely scenario however due to the presence of the ~180° chain inversions.

Sample	Processing program	Distance between subunits (Å)	Number of particles	Angle between subunits (°)
Heat polymer + 4B12 Fab	IMAGIC	64	6,036	16
Heat polymer + 4B12 Fab	Relion	67	6,923	59
Liver polymer + 4B12 Fab[250]	IMAGIC	73	3,749	11
Liver polymer + 4B12 Fab[250]	IMAGIC	68	4,371	16
Liver polymer + 4B12 Fab	Relion	73	6,848	88
Liver polymer + 4B12 Fab	Relion	69	9,239	60

Table 5.3. Summary of the dimer models generated from the two image-processing softwares for two different samples. *The liver polymer + 4B12 Fab, processed using IMAGIC, was carried out by Dr. S. Faull[250].*

Multiple attempts were made to identify other dimer configurations embodied within both datasets, however these were unsuccessful. It is clear that other conformations exist, for example the big-angle particles identified from the heat polymer IMAGIC processing, however these are likely to represent minor species where too few images have been identified to generate a sensible reconstruction. It is also possible that the 'stump' models generated in the Relion processing workflows represent real dimer structures, where only one out of the two AAT subunits are labelled with Fab fragment. Most class averages in a 2D classification of the particles making up the 'stump' structures show one subunit without any protruding Fab density (Figure 5.15A). However, generally this subunit appears to be brighter, suggesting the Fab is indeed bound and is orthogonal to the plane of the micrograph. The presence of Fab is further supported by the fact that polymers were incubated with excess Fab, the interaction between 4B12 and AAT is stable during purification and the lack of

significant gaps in labelling on inspection of micrographs (Figure 6.12C). Even if the dimer was part unlabelled, it would be difficult to obtain useful information regarding the rotational relationship between subunits at this resolution.

5.4.2 The C-terminal model is most consistent with the reconstructions

The linkage between subunits in the pathological AAT polymer has been a subject of debate for more than 25 years. Out of a total of five models that have been proposed, the loop-β sheet A, β-hairpin and C-terminal models are deemed most likely to represent the polymer structure[5,32,78,83-86]. Linear versions of these models were created previously and used to test their fit into the two IMAGIC *ex vivo* polymers generated by Dr S. Faull's work[250]. It was found that the C-terminal model was most consistent with both structures and the loop-β sheet A model with the minimal insertion was possible but inconsistent[250]. Similar conclusions were found in this study when the three polymerisation models (with the refined AAT/Fab interactions) were fit into the Relion *ex vivo* 60° and 88° models, the Relion heat 59° model and the IMAGIC heat 16° model. The differing inter-subunit angles allowed an assessment of how well the polymerisation models could accommodate such a structure. Therefore, the additional *ex vivo* NS-EM structures and the heat NS-EM structures have enabled a more thorough analysis of the fit of the polymer models, contributing towards resolving the controversy.

The distance and angles between the subunits in the reconstructions meant that the linker of the loop-β sheet A model pulled the adjacent AAT subunit out of optimum positioning in the density. The loop-β sheet A model involves the exchange of the RCL between neighbouring subunits[32]. This was largely based on the ability of RCL peptides to block polymerisation but there is no structure for such a linkage and so the exact residues that are exchanged has not been explicitly defined[32]. It is currently proposed to be residues P8-P3 that swap and accordingly, this was modelled in this study[73]. The results of the fitting, combined with assertions that propose this linkage would not produce the known stability or flexibility of polymers, suggests that the loop-β sheet A model is unlikely to represent the linkage in heat or liver polymers[91]. Although RCL insertions to different extents could be tested, i) a longer RCL swap would further limit flexibility of the chain and ii) a shorter RCL swap could not justify the extreme stability of polymers to denaturant and heat.

The β-hairpin model, modified from that originally proposed[78] by retaining an intact helix I, was adequately accommodated by the densities and thus on this data cannot be discounted as the linkage[78]. The model involves the exchange of the RCL and strand 5A; this linkage was observed in a crystal structure of a closed antithrombin dimer[78]. The dimer was formed by incubation at low pH and experiments in support of the structure, used denaturants to induce polymerisation. Considering 2C1 reactivity towards heat and liver polymers, but not denaturant and acid polymers, the β-hairpin model seems unlikely to be relevant to Z-AAT liver polymers[69,76], although there is evidence it may be applicable to at least one rare variant[191].

Fitting of the C-terminal model resulted in the highest correspondence between structure and density. This model describes the exchange of strands 1C, 4B and 5B between subunits and the self-insertion of the RCL[86]. A crystal structure of a closed AAT trimer, which was formed by heating and consequently recognised by 2C1, shows this linkage. It is the most flexible model – in keeping with the observed flexible nature of AAT chains – but the crystallised species was an atypical, disulphide-trapped form and its relevance to natural polymers was not established[76,86]. However, considering all the published data and the data presented here, the C-terminal model appears to be the most plausible out of the proposed models.

As both the liver polymer and heat polymer structures were best explained by the C-terminal model, this reinforced their structural similarity in addition to the comparisons described earlier. At the resolutions of the structures, the fitting results are only indicative; a definitive answer would require high-resolution cryo-EM data. It is possible that AAT subunits form polymers through an alternate mechanism, different from those presented in the models, but such a proposal could not be made with these data. Furthermore, liver AAT polymers resulting from different mutations may not necessarily have the same structure. The polymers in this study were isolated from a ZZ-AAT homozygote, and previously 2C1 has been shown to recognise AAT polymers resulting from a range of mutations including the predominant Z mutation and rarer King's, M_{pisa} and $E_{taurisano}$ mutations[76,260]. However, recently pathological polymers formed by the Trento mutation showed reduced recognition by 2C1 indicating a possible structural difference[191]. This means conclusions drawn from studies of ZZ-AAT polymers may not necessarily extrapolate to polymers formed by other mutations. On a similar note, it is generally believed that

pathological polymers of other serpins form by the same mechanism[41]. This would also need to be addressed as serpins have RCLs of varying length and a longer RCL would result in the loop-β sheet A model becoming more permissive.

5.4.3 Improvements incorporated into the processing of the heat polymers compared to the *ex vivo* polymers using IMAGIC

Image processing of the *ex vivo* AAT polymers using IMAGIC resulted in density that was very low resolution (around 30Å) and that failed to show the distinctive contoured density associated with the Ig domain. Accordingly, improvements to the workflow were incorporated for the heat-induced polymer. For the heat-induced polymer, linear regions of polymer were selected from the micrographs using the helical function of EMAN/Boxer[254], providing a more efficient way to pick particles compared with the manual selection performed in the prior study. This introduced less bias into the selection process where previously only discernible dimers were picked. It also increased the dataset size, allowing a more robust classification and selection step at later stages of processing.

Eigenimage analysis was used to guide the division of particles into subsets for the processing of both datasets. However, for the liver processing, the data were divided into subsets representing different distances between subunits. This difference was reflected in the major eigenimage, yet as the alignment of images was carried out against a reference with two subunits, it is likely this did not represent two populations that varied in inter-subunit distance. For this to be true, the eigenimage would have had to result after particle alignment to a reference of a single subunit. The difference in distance between subunits that was observed could be due to a different tilt of the projections and so the particles were classified by view rather than population. It is possible that the particles were over-classified and they are all consistent with a small-angle dimer reconstruction with an angle around 11° or 16°. Instead in the heat polymer processing, images were divided based on eigenimages representing the clear heterogeneity observed between the relative rotation of subunits due to the positioning of the Fab fragments. Only once this division had happened, did a sensible dimer model with density for both Fab fragments result indicating homogeneity of particles is important in obtaining reconstructions[129].

5.4.4 Comparison between image processing of AAT polymers using IMAGIC and Relion

IMAGIC and Relion are two image processing software packages for EM data[121,122,162,259]. They both align images in 2D and produce 3D reconstructions but are based on different algorithms and statistical methods. Image processing was initially carried out for the heat-induced polymer and liver polymer datasets in IMAGIC and then repeated in Relion.

A major difference between the workflows was the number of particles used in image processing. Over eight times more particles and twice as many were used in the Relion processing of the liver and heat polymers respectively, in comparison to the IMAGIC processing. This increase in particle number was a result of automated particle picking, however theoretically autopicked particles could have been inputted into the IMAGIC software. The IMAGIC particles used were manually picked using EMAN[254]. Obvious stretches of polymer were selected which was both time-consuming and tedious but reasonably accurate. Manual picking may have excluded rare but important views, limiting the reconstruction resolution and accuracy. In comparison, autopicking is much quicker and selects more particles. A certain amount of 'junk' particles are expected in autopicking however these are easily removed through rounds of 2D classification. It is also important to low-pass filter autopicking references to avoid picking noise[141-143]; for the Relion processing, autopicking references were filtered to 30Å to avoid this bias.

In both processing pipelines, particle images were selected at the level of 2D and 3D. In IMAGIC, eigenimages analysis was useful in displaying heterogeneity between 2D particles and this directed the division of the particles into small-angle and big-angle subsets. Images were also selected based on 3D-error scores at the 3D reconstruction stage. In Relion, particle heterogeneity was sorted by 2D classification and 3D classification jobs. An important consideration when comparing the processing carried out for the AAT polymers is that the reconstructions obtained in IMAGIC were not replicated in Relion, either within a full processing pipeline or with a more limited re-appraisal of the identical particles and reconstructions from IMAGIC. Both software package utilise different approaches: IMAGIC implements MSA and eigenimages are used to show variations between images[121] whereas Relion is based on ML methods[259]. This outcome was primarily addressed in the Relion

processing, where multiple attempts to reconstruct a small-angle dimer (like those generated from the IMAGIC processing) were made but all were unsuccessful. These included using the small-angle dimer as a reference for 3D classification to help selection of relevant particles and using the exact particles that constituted the IMAGIC reconstructions. It raises the question about the reliability of the IMAGIC models, especially considering the class averages of the liver polymer IMAGIC particles seem to resemble a dimer with an angle greater than 11-16° between the subunits (Figure 5.15B). This may represent limitations of the fundamental basis for particle reconstruction employed by IMAGIC, or the higher degree of subjective user choice in the processing pipeline; one of the stated intentions of the Relion developers is that it requires less decisions from the user that can impact the processing outcome[122]. For example, alignment of the tetramers against a rectangular stripe, with 2D classification according to specified eigenimages into classes with relatively few equally weighted members, would explicitly exclude tilt-views or end-views, which would not be interpretable according to the decision tree used. While there is also a subjective decision process in choosing to progress or not a given 2D or 3D class in Relion, these classes are far less focussed, instead incorporating a variable number of images which are weighted according to a calculated probability score. It is clear that one by-product of this approach is that there are a proportion of particles that are 'misclassified' according to the shape represented by the total sum, but bear statistical features warranting their inclusion. Furthermore, the explicit treatment of heterogeneity favours inclusivity of classes by the user with reduced risk of influencing the outcome, if a 2D class is incorrectly selected.

In general, the reconstructions generated by Relion had much better subunit definition and were solved to a higher resolution. The IMAGIC models were solved to a resolution of between 30-33Å at an FSC of 0.5 whereas the Relion models were solved to a resolution between 23-27Å. The highest resolution model was the liver 60° dimer, probably explained by being the reconstruction with the most constituent particles (Figure 5.14). The subunits in the Relion models were better defined at the interface between AAT and the 4B12 Fab and displayed depressions in the density where the hole of the Fab structure was modelled. These subunits enabled a better model of the interaction between AAT and 4B12 Fab to be created, which was subsequently used to assess the fits of the polymerisation models into the reconstructions. The modelling of the single subunit gave conclusive results in the

positioning of the heavy and light chains of the Fab; this demonstrates the relatively high resolution achieved in this study using NS-EM.

5.4.5 Addressing concerns about model bias

Orientation determination of particles by projection matching requires a reference model. An incorrect starting model can bias or invalidate projection matching[129], producing an incorrect reconstruction. It is difficult to tell the extent of reference bias on the final model. Therefore, it is important to use a reliable starting model, low-pass filter the reference to remove surface features and validate structures by other biochemical and biophysical experiments[129].

The processing workflow for each of the reconstructions discussed differed in the starting model for projection matching. The approach used was gradually adapted to reduce potential model bias: the starting model progressed from being a dimer model, to a 'stump' model and finally to a monomer model. This change reduced the influence of bias away from a specific angle of rotation and then from a specific distance between two subunits, finally to a reference unbiased for the angular relationship within a dimer. The starting models were always low-pass filtered to greater than, or equal to, 30Å to remove any higher-resolution features that could bias projection matching.

In the IMAGIC processing of the heat-induced polymer, one liver dimer reconstruction, which had an 11° angle between subunits, was used as an initial model. Projection matching was iterated several times until convergence on a model with a 16° angle between subunits was achieved. The change in angle, and convergence on an angle, suggested that that the reconstruction was moving away from the initial model and was converging towards the model inherent in the experimental data. This was supported when projection matching was carried out against reference models constructed with angles between 0-60°. All the resulting reconstructions had inter-subunit angles closer to 16° than the reference model used for orientation determination. Although addressing model bias was not completed to the point of convergence on 16°, it does support the presence of a small-angle dimer within the data. Furthermore, there was a good match between the projections and the reprojections of the 16° reconstruction indicating reliability of the structure. In hindsight, it would have been best to not use a liver dimer model as a reference in

the heat-induced polymer processing for a study comparing the heat-induced and liver polymer dimer structures to limit the influence of model bias. However, this was the only dimer model available based on experimental data. A possible approach that could have been attempted given time would have been to try a monomer reference, or if this did not work the 'stump' reference, to allow the second subunit and hence rotation between subunits to purely come from the data. However, one of the likely sources of failed reconstructions using these 'reduced' models in Relion is that the AAT-Fab subunit is approximately the same length and shape as the AAT-AAT subunit at these resolutions (Figure 5.18). It may be therefore that there was an increased mis-assignment of Euler angles to some of the projections, confounding the reconstruction. An alternative approach, considered best practice, is to produce an *ab initio* model using the common line theorem to assign orientations to projections or from an experimental technique that directly provides orientations such as random conical tilt reconstruction (2.9.5)[129]. However, such approaches are difficult for small, heterogeneous and asymmetric samples and the flexible dimers considered here represent very poor subjects for such an analysis.

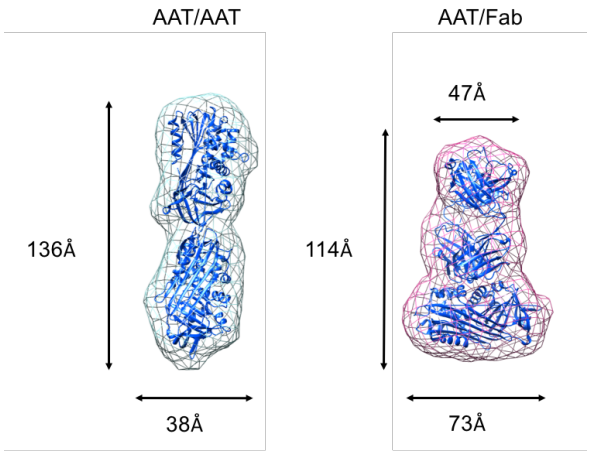


Figure 5.18. Comparison of the dimensions and shapes of AAT/AAT and AAT/Fab. *Electron density at 20Å was simulated from the atomic structures and measurements were made in Chimera.*

The ease of image processing with Relion facilitated more analysis about the influence of model bias and allowed more checks to be carried out. The heat polymer processing involved the 60° liver dimer as a reference for 3D classification, which produced a model very similar to the reference and required caution regarding model bias. The particles selected from this model however were refined against a single

subunit model, and the second subunit of the 59° dimer was entirely reconstructed from the data. This provided confidence that while the choice of reference might have enriched for components compatible with this angle, it did not dictate the consensus reconstruction of the final structure, which includes alignment and weighting of the contributions of individual components. Processing of the liver dimer used several rounds of 3D classification from the final 68,000 particles to select images. All these steps used either the stump model or a single subunit model as a reference. Therefore, the resulting 60° and 88° models were least influenced by references regarding their inter-subunit rotation. These models also reappeared in the search for extra reconstructions using a monomer reference highlighting their dominance in the dataset. In the early stages of the liver polymer and in the heat-induced polymer processing, a single subunit was used as a reference, however poorly defined models resulted. This indicated that if particles were sub-optimal (for example included junk particles or have low SNR) then projection matching struggled.

Relion also provided the opportunity to obtain an *ab initio* reconstruction. Although this function is usually used to generate a starting model for projection matching that will later be refined, here it served a useful purpose to determine the influence of potential model bias. The images corresponding to the 59° heat-induced, 60° liver and 88° liver dimer all produced models comparable to the final refined reconstructions, providing further confidence of their existence in the data.

Model bias was particularly apparent in the searches for extra dimer reconstructions within Relion. When the 16° IMAGIC heat-induced and 88° Relion liver dimers were used as references in the 3D classification of the heat polymer particles, a model similar to the reference was generated. However, the corresponding particles did not refine against a monomer, unlike the 59° Relion heat-induced dimer. Similar results were apparent when searching for other dimer structures in the liver dataset. Therefore, the model produced after 3D classification was strongly influenced by the reference. On the other hand, particles could not be forced to even refine against a dimer reference model. This most likely reflects the extra safeguards against bias present in the refinement algorithm as opposed to the 3D classification algorithm[141].

5.4.6 Monoclonal antibodies aided EM studies

In this study, a mAb has been used to aid visualisation and angular assignment of AAT polymers during image processing. This is a variation on the frequent use of antibodies in EM as an immunolabelling reagent to determine identity or location of a component within the electron density of a complex. For example, a gold-conjugated antibody has been used to confirm that polymers in a partially purified sample were formed by AAT[3]. Similarly, comparisons between antibody-bound and non-bound samples have been useful in determining location: outer capsid proteins were localised on bacteriophage capsids by analysing micrographs[261] and a specific protein in a much larger secretion system complex was localised from the extra electron density apparent in class averages[18]. Antibody fragments are becoming a useful tool in cryo-EM studies of small proteins, helping to increase the effective mass of the protein and facilitating image alignment or stabilising certain conformational states[262,263]. The former is the reason why Fabs have been employed for cryo-EM studies of the AAT polymer in Chapter 6. Antibody fragments have also been employed in X-ray crystallography to help form favourable crystal contacts for proteins resistant to crystallisation[264]. A common concern when using antibodies in structural biology is whether antibody binding affects the structure of the protein being studied. In this case, it is unlikely that the binding of 4B12 Fab significantly alters polymer structure. There may be small conformational changes at the site of binding or even sites more distant[229,235] however it was deemed unlikely to alter the region of interest – specifically, the linkage between subunits – in pre-formed polymers. 4B12 does not structurally discriminate between AAT species and polymers are so stable that they require 8M urea or heating to greater than 100°C to dissociate[12,67,116,117].

Nevertheless, image processing of the undecorated heat polymer was undertaken to address the possibility of antibody-induced structural changes in the polymer. A NS-EM dataset was collected of this sample, however as expected there were difficulties with angular assignment and identifying the rotation between subunits (not shown). This was probably due to the featureless nature of AAT at low-resolution. The difficulties discussed, together with the finding that it was not possible to reconstruct the Relion 59° heat-induced dimer when using a reference lacking the Fab domains, highlights the importance of using the 4B12 Fab to label AAT polymers in EM structural determination. A more feasible approach to understand the interaction

between 4B12 Fab and AAT would be to solve a crystal structure of the complex. Obtaining diffracting crystals of this complex has proven elusive, and crystallisation trials are ongoing. A complex structure could also improve the accuracy of the fitting of the polymerisation models into the NS-EM reconstructions as currently the 4B12 Fab-AAT interaction is a model based on epitope mapping data[117]. However, obtaining a crystal structure is not assured – it is a process dependent on the ability to form a regular repeating lattice, and sample heterogeneity and flexibility are confounding factors that can interfere with this process. Furthermore, a constrained crystal 'snapshot' would not represent any conformational mobility observable by EM.

5.4.7 Conclusions and future directions

This work has demonstrated that the structural comparison of heat-induced and *ex vivo* polymers by NS-EM is both feasible and informative. Low-resolution reconstructions were made of several dimer species, where each individual subunit in the Relion reconstructions had high enough definition to improve the modelling of the 4B12 Fab-AAT interaction. Generally speaking, the two polymer populations have components that can be said to be structurally similar at this resolution, in keeping with the common recognition by 2C1[76]. Given the assumption that each sample contains a single polymer configuration, the observation of a directly equivalent structural component makes it likely that those configurations are the same.

From a comparison of the loop-β sheet A, β-hairpin and C-terminal models of polymerisation, it was the latter that best fit the structures[86]. NS-EM can provide information on molecule shapes but the internal details are not revealed; at the resolution obtained, approximately 20Å and 30Å from the Relion and IMAGIC reconstructions, respectively, this correspondence is informative but not definitive. An assessment of the internal structural details of the polymers requires reconstruction from cryo-EM data. The cryo-EM structure of *ex vivo* polymers is addressed in Chapter 6. A further consideration regarding differences between NS-EM and cryo-EM is the degree to which the sample is altered by the fixation technique used. Negative staining can cause structural artefacts such as distortion or flattening of structures due to the deposition of heavy atoms[129,265]. The properties of the stain can also affect the sample for example uranyl acetate has a low pH, which could cause sensitive proteins or complexes to disassemble, and it binds to lipid and nucleic acid

phosphate groups, which could induce sample aggregation[266]. Cryogenic freezing of samples yields a more native and hydrated structure, although particles present in very thin sections of ice, may still be structurally distorted if they are not fully embedded in the ice[267].

Of importance to the research into the mechanism of polymerisation of serpins, is the distinction between the different properties of polymers that are induced under different *in vitro* conditions (heating-, denaturant- or acid-induced). Different studies in the literature have used polymers made under different conditions to infer details of polymers and the mechanisms by which they polymerise[32,69,72,75,92,268]. Obtaining NS-EM structures for other artificial polymers labelled with 4B12 Fab, for example those induced by urea or acid, would now help inform *in vitro* studies of AAT polymerisation. The structural differences proposed on the basis on 2C1 immunorecognition may be explained and published experiments could be re-interpreted[69]. Preliminary NS-EM micrographs show that urea or acid-induced polymers are shorter and subunits are less bead-like. Furthermore, during this work, another lab member found a novel method to induce polymerisation artificially by incubation with 20% w/v MPD. MPD polymers were recognised by 2C1 and had similar appearance to heat-induced polymers by NS-EM and so these polymers could also be structurally characterised in more detail. An alternative method to heating – such as MPD incubation – that replicates the linkage in liver polymers, would be useful from a practical standpoint as it would allow the monitoring of changes in a sample without the need for accurate temperature control.

6.0 Cryo-EM of *ex vivo* AAT polymers

6.1 Introduction

6.1.1 AAT accumulation in the liver

Misfolded and polymerised Z-AAT is retained within hepatocytes as ER-derived intracellular inclusions. Inclusion bodies (IBs) were first identified around 1970 in autopsy studies of AATD patients[269,270]. Individuals had been diagnosed with early-onset emphysema due to a circulating deficiency of AAT, but these studies showed they had also developed cirrhosis and suggested a common hereditary cause. The IBs from liver samples of individuals with AATD were then characterised and found to be rounded, single membrane-enclosed and of 1-20 μm in diameter[271]. They were periodic acid-Schiff (PAS)-positive indicating the presence of glycans and immunofluorescence studies showed they contained accumulated AAT[271]. This appeared to confirm that the plasma deficiency of patients resulted from accumulation rather than failure of AAT production; later it was found that degradation was also a major contributor. Although typically IBs are thought to contribute to liver cell toxicity due to the association with liver disease, it has been postulated that IB formation is a protective mechanism of the cell that maintains a functional secretory pathway[105].

6.1.2 Cryo-EM of *ex vivo* protein aggregates

The difficulty in obtaining and/or working with small amounts of patient-derived tissue has meant many diseases have been studied by establishing *in vitro* models. The concern, however, is that there are structural differences between *in vivo* material and *in vitro* assembled molecules. This potential distinction is an important question regarding the structure of AAT polymers, and is addressed throughout this thesis.

Questions regarding the comparability of *in vivo* and *in vitro* models have similarly been raised within Alzheimer's disease research. In Alzheimer's disease, for example, filaments of tau protein formed within human brains are much more effective than *in vitro* assembled filaments at inducing disease in mouse models[272]. Cryo-EM structures of tau filaments purified from the brain of a patient diagnosed with Alzheimer's disease have recently been determined to 3.5Å resolution[272]. The

structures reveal differential lateral packing between protofilaments in paired helical and straight tau filaments. In contrast to AAT polymers, which are extremely flexible, tau filaments are helical and rigid, enabling the use of a helical reconstruction algorithm (such as that available in Relion). An alternative fold of tau filament has also recently been solved to 3.2Å resolution by cryo-EM[273]. These filaments were isolated from a patient with Pick's disease, a neurodegenerative disorder characterised by frontotemporal dementia. These studies of filaments from patient samples are key to revealing how tau can adopt distinct folds in different neurodegenerative diseases (Figure 6.1); it is hypothesised that conformers of tau filaments give rise to the clinical phenotype that defines each disease[272,273].

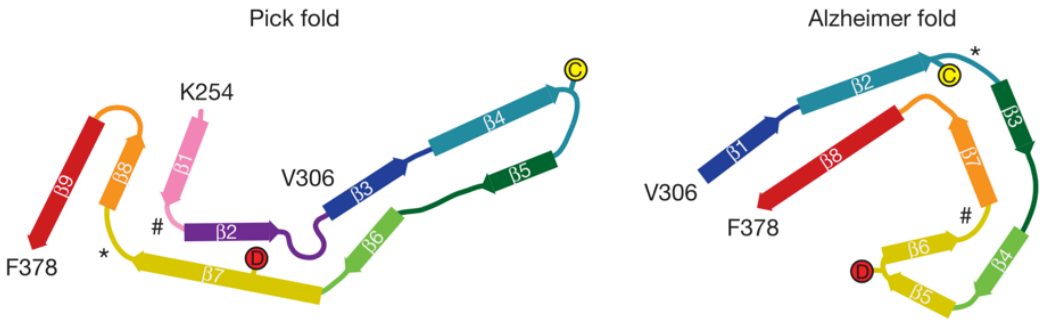


Figure 6.1. Schematic of the folds of Tau protein in Pick's disease and Alzheimer's disease. *The structure of the fold was determined from filaments isolated from brain samples of such patients. Figure reproduced from[273].*

The aggregation of α-synuclein, which underlies neuropathological diseases such as dementia with Lewy bodies and Parkinson's, has also been studied using cryo-EM. Fibrils were determined to 3.4Å resolution and revealed two composite protofilaments with a step-like β-sheet arrangement[274]. This structure was of human recombinant α-synuclein but allows hypotheses about the mechanism of fibril destabilisation by common mutations associated with Parkinson's disease to be made. The authors postulate that a different fibril structure could be formed from mutated α-synuclein, highlighting the importance of studying disease mutations or patient samples.

Lastly, advanced cryo electron tomography methods have allowed visualisation of aggregates of polyglutamine inclusions, like those formed in Huntington's disease, in unperturbed cellular contexts. IBs of intact, transfected neurons from mouse brains, as well as transfected human cells, were found to consist of amyloid-like fibrils[275].

The polyQ fibrils interacted with cellular membranes, in particular deforming the ER membrane[275,276]. These findings led the authors to hypothesise that the formation of fibrillar IBs (in addition to the toxic aggregates) contribute to cell pathology, rather than having a protective role[275,276]. Equivalent studies into AAT IBs may be useful in gaining insight into the role of IBs in liver pathology and the *in situ* structure of AAT aggregates inside cells.

6.1.3 Aims of the AAT cryo-EM work

The resolution of a 3D reconstruction that can be achieved using NS-EM images is limited by the grain size of the metal stain. The highest resolution achieved using NS-EM is of a helicase at 18Å and this sample had the advantage of possessing rotational symmetry to aid in the reconstruction[277]. Cryo-EM under optimal conditions has the potential to achieve near atomic-resolution and now regularly achieves 5-8Å resolution, enough to resolve α-helices. Recently, resolutions of 2.2Å and 2.9Å have been obtained from structurally homogenous proteins, with thorough optimisation of all stages of the process[278,279].

This part of the project aimed to obtain a high-to-medium resolution structure of the *ex vivo* polymer using cryo-EM. Polymers were extracted from the liver tissue of a ZZ homozygote patient, characterised biochemically and prepared for cryo-EM studies. A structure would hopefully provide information on the nature and extent of the intermolecular interaction between subunits within the AAT polymer.

6.2 Methods

6.2.1 Purification and characterisation of *ex vivo* polymers

6.2.1.1 Extraction of *ex vivo* AAT polymers from liver samples

Liver samples were obtained from the University of Birmingham BioBank. AAT polymers were extracted from liver samples of an ZZ AAT homozygote who had undergone an orthotropic liver transplant, using a protocol adapted from previous polymer extractions[280,281]. Frozen blocks of liver tissue, totalling approximately 7g, were thinly sliced and incubated at 37°C for 1 hour with rocking in 10ml of Hank's balanced salt solution (modified with sodium bicarbonate without phenol red, calcium chloride and magnesium sulphate, Sigma-Aldrich) with 5 mg collagenase (type 1A, from Clostridium histolyticum, Sigma-Aldrich). Fibrous tissue was removed from the resultant suspension by filtering through BioPrepNylon synthetic cheesecloth with 50 µm pore size (VWR). The filtrate was centrifuged at 3000 g at 4°C for 15 minutes. The pellet was resuspended in 3 ml of 0.25 M sucrose in 5mM EDTA, 50mM NaCl, 50mM Tris pH 7.4 buffer that had been filtered, by sonicating on ice for 45 seconds at 2.5 microns using a soniprep 150 sonicator (MSE). The sample was layered onto the top of two 14ml 14x95mm centrifuge tubes (Beckman Coulter) containing a 0.3M-1.3M sucrose gradient made up in buffer. The gradient was made using a Gradient Master Model 108 (BioComp) and was kept at 4°C for an hour before applying the sample. The sample was centrifuged at 25000 g for 2 hours at 4°C in an Optima L-90K Ultracentrifuge (Beckman Coulter) with a SW 40 Ti rotor. The supernatant was discarded and the pelleted inclusion bodies were snap frozen using liquid nitrogen for storage overnight at -80°C.

Half of the inclusion bodies were processed per preparation and transferred from the ultrafuge tube to a microcentrifuge tube. The pellet was resuspended and washed 5 times with 500 µl of 5mM EDTA, 50mM NaCl, 50mM Tris pH 7.4 buffer by centrifuging at 16,000 g for 15 min at 4°C. After washing, the inclusion bodies were resuspended in fresh buffer and sonicated on ice for a total process time of 6 min of 15 sec on /15 sec off pulses at 2.5 microns to release soluble polymer. The sample was centrifuged at 16,000 g for 15 min at 4°C and the supernatant retained and moved to a fresh microcentrifuge tube. This step was repeated until all visible insoluble material was removed.

6.2.1.2 Purification of *ex vivo* AAT polymers

The supernatant released after sonication was diluted in 20mM Tris pH 8 prior to loading onto a pre-equilibrated 1 ml Hi Trap Q HP column (GE Healthcare). Protein was eluted by a 0-100% gradient of 20mM Tris, 1M NaCl, pH 8 over 20ml and fractions analysed by SDS-PAGE to identify the fractions containing AAT.

6.2.1.3 Purification of *ex vivo* AAT polymers labelled with Fab (development from method 3.5.4.2)

After incubation of *ex vivo* polymers with a three-fold molar excess of 4B12 Fab for 3 hours, the sample was diluted into 20mM Tris pH 8 buffer. The sample was loaded and eluted from a 1ml Hi Trap Q HP column (GE Healthcare) as in 6.2.1.2.

6.2.1.4 Deglycosylation of AAT

AAT was deglycosylated by PNGase F and EndoH (NewEngland Biolabs) according to manufacturer's instructions. A solution containing 20µg of AAT and 1µl of glycoprotein denaturing buffer made up to 10µl with water was denatured at 100°C for 10 min. The solution was briefly chilled on ice and centrifuged before adding 2µl Glycobuffer (buffer 2 for PNGaseF and buffer 3 for EndoH), 1µl PNGase F or Endo H and water up to 20µl total volume. 2µl 10% NP-40 was added for the PNGase reaction as PNGaseF is inhibited by SDS. The reaction was incubated at 37°C for 2 hours and run on a SDS-PAGE. The negative control contained no deglycosylating enzymes.

6.2.2 Preparation of samples for screening antibody strategies to reduce flexibility

6.2.2.1 Double-labelling of heat polymers with Fab

Three-fold molar excess of both 9C5 Fab and 4B12 Fab were incubated with heat-induced polymer in PBS for 3 hours at room temperature. The sample was diluted to 0.04 mg/ml in Tris-based buffer prior to application on grids for NS-EM imaging according to method 3.5.4.4.

6.2.2.2 Labelling of heat-polymers with Fab_2 or mAb

4B12 Fab_2 was made using the Pierce Mouse IgG1 Fab and $F(ab')_2$ Preparation Kit (Thermo Scientific) according to the manufacturer's instructions. Heat polymers were incubated at a 1:0.5 molar ratio with either 4B12 Fab_2 or mAb for 3 hours at room temperature and prepared for NS-EM as in 6.2.2.1.

6.2.3 Cryo-EM methods

6.2.3.1 Freezing of grids

The Vitrobot was pre-equilibrated to 4°C and 95-100% humidity for 30 min before freezing. Ethane gas was condensed into a copper vial, in the centre of a container pre-cooled to liquid nitrogen temperatures. The freezing process was started when a thin layer of solid ethane formed at the bottom of the vial. Lacey Carbon Films on 300 Mesh Copper grids (Agar scientific), Quantifoil R 1.2/1.3 or UltrAuFoil R 1.2/1.3 Holey Gold films on a 300 Gold mesh (Quantifoil) were glow-discharged at 30 mA for 1 min. 3µl of purified sample at 0.7 mg/ml or 0.5 mg/ml in EM buffer was resuspended, applied and incubated for 30 sec before blotting once for 3-5 sec with a blot force 5, 10 or 15 and drain time of 0. The sample was then rapidly plunged into liquid ethane and transferred into a grid box, maintaining the grid at liquid nitrogen temperatures.

6.2.3.2 Screening of grids

Cryo grids were screened using a Tecnai 200 kV FEG with a Direct Electron (DE-20) direct detector. The sample was imaged at a nominal magnification of 29,000X, with a pixel size of 1.54Å, at a defocus 1-4 µm using serialEM[282].

6.2.3.3 Data collection

A dataset was collected on a FEI Tecnai G2 Polara with a 300kV FEG and a K2 direct detector (Gatan). Areas suitable for data collection were first identified at low magnification by capturing a low magnification montage map (LMMM) followed by a medium magnification montage map (MMM). Images were collected using automated data collection where polygons of suitable holes were defined. A total of 221 micrographs were collected at a nominal magnification of 200,000X and a

defocus range of 1-4 μm. The pixel size for images was 1.1Å. Each micrograph was taken over a 12 second exposure with 60 subframes captured every 0.2 sec. The dose per frame was 0.95 electrons per $Å^2$ making a total dose of 57 electrons per $Å^2$ for the full 60 frames.

6.2.3.4 Micrograph preparation

Image processing was carried out within Relion v2.1[141]. Micrograph movies were motion-corrected using Motioncor2 from frame 1 to 60 with dose-weighting for initial processing and later, between frames 2 to 25 for processing of particles exposed to less electron damage. Defocus values were estimated between 0.5-5 μm for the aligned micrographs using CTFFIND-4.1. CTF fits were manually inspected and a total of 217 micrographs kept for further processing based on adequate CTF correction.

6.2.3.5 Processing for reconstruction of the dimer

6.2.3.5.1 Producing an initial cryo-EM dimer model

Around 1000 dimer particles were manually picked from the micrographs and extracted into a 180-pixel (198Å) box from micrographs that had been motion-corrected from frames 1-60. The extracted particles were subjected to 2D classification with a 196Å mask into 10 classes. Well-defined class averages, representing different views of the dimer, were selected and used as references for autopicking. Autopicking parameters were first optimised on four representative micrographs. The settings chosen for autopicking of the entire dataset were: a 60Å distance between particles, a 0.8 picking threshold, 1.1 background standard deviation and 40Å low-pass filter of the references. Around 103,000 particles were extracted for 2D classification into 100 classes from 2-25 frame motion-corrected micrographs. Classes (and corresponding particles) were selected for three further rounds of 2D classification and selection to obtain a more focussed dataset of 16,000 particles. These particles were subjected to 3D classification into four classes using the 16° IMAGIC NS-EM model filtered to 50Å as an initial model. The resulting model with the closest resemblance to the expected appearance of the dimer had one well-formed subunit. This was copied and duplicated onto the least-formed subunit using

Chimera[9]. This new initial model was used as a reference for another 3D classification. The best resulting model had two subunits of AAT/Fab shape.

6.2.3.5.2 Refining the cryo-EM dimer model and a monomer model

The new cryo-model was reprojected in a uniform angular distribution. Projections clearly depicting two subunits were selected as references for autopicking with parameters of 60Å distance between particles, a 0.5 picking threshold and 1.1 background standard deviation. Around 247,000 extracted particles were reduced to 108,000 good particles by three rounds of 2D classification and subset selection. Class averages were picked that visually resembled dimers with a 'small-angle' between subunits, leaving 93,000 particles. These particles were used as input into a 3D classification to generate four classes. Particles corresponding to models with at least one well-formed subunit were selected (models 1 and 3). The process of 3D classification was repeated.

Particles were grouped according to whether just one subunit was well-formed (23,000 particles, models 2 and 4) or whether a second subunit was appearing (15,000 particles, model 3). Further 3D classification of these groups did not produce models with differing characteristics. The 15,000 dimer particles were refined separately against the autopicking cryo dimer model and a single subunit reference. The resulting structures were analysed in Chimera[9]. Distance and angles between subunits were measured as in 5.2.2.5. A mask was made in Chimera by thresholding at a volume calculated from the molecular weight using an estimated density of 1.2Å^3/Da, erasing noise, applying a 10 standard deviation Gaussian filter and thresholding again. The mask was imported into Relion and a soft-edge of 20 pixels applied. The mask and half-map from the refine job was used in a post-process job to determine resolution by Fourier Shell Correlation (FSC). A similar approach was used for the 23,000 particles: they were refined against a single subunit reference, with a relaxed mask around the single subunit and subjected to post-processing for resolution estimation.

6.2.3.6 Processing for reconstruction of the monomer

The monomer reconstruction from the dimer processing pipeline was reprojected using the Relion_project command between 30-150° on the Euler sphere. These

served as references for autopicking using the parameters: 60Å distance between particles, a 0.5 picking threshold and 1.1 background standard deviation. Around 253,000 particles were extracted in a 144-pixel box and sorted by four rounds of 2D classification and subset selection with 136Å circular masking to reach 63,000 particles. The resolution limit in the E-step of the 2D classification was limited to 8Å and the background set to either zeros or random noise. Particles were further selected based on reconstructions after 3D classification. 3D refinement was carried out both on particles selected after 2D classification and 3D classification. A mask was made in Chimera by applying a threshold at a volume corresponding to 100 kDa molecular weight, erasing noise, applying a 10 standard deviation Gaussian filter and thresholding again. The mask was imported into Relion and a soft-edge of 20 pixels applied. The final resolution was calculated using the FSC as estimated by interpolating the FSC curves at values of 0.5 and 0.143, as both conventions are in routine use. Attempts to increase the reconstruction resolution by improving the alignments included altering the range of the alignment search (from 5 to 10 pixels), padding the images to 300 pixels, altering the T value between 1 and 4 and masking in 3D and 2D.

6.3 Results

6.3.1 Extraction of *ex vivo* AAT polymers

To enable structural and biophysical characterisation of the *ex vivo* AAT polymer, polymers were extracted from liver samples of patients with AATD. These patients had undergone orthotopic liver transplantations due to advanced liver failure associated with the accumulation of AAT polymers in IBs. The method used for polymer extraction was based on a published method but modified to avoid the use of detergents and denaturing reagents. This was to avoid any possible unwanted structural changes to the polymer structure as such reagents are known to induce polymerisation or depolymerisation. Instead, limited sonication in successive rounds was used to break open the IBs, releasing the soluble AAT polymer. Sonication at the amplitude and for the duration used has been shown to not induce depolymerisation and enabled isolation of AAT *ex vivo* polymers for NS-EM in work conducted previously (Dr S. Faull, personal communication).

Polymers were extracted from a liver sample taken from a 46-year-old ZZ homozygote using the method summarised in Figure 6.2. The supernatant (containing the polymers) was run on an SDS-PAGE gel and showed that although there were many contaminating proteins still present, there was a dominant band present that was around the molecular size of AAT monomer at 52 kDa. One other band showing enrichment over the others (marked by an * in Figure 6.2) was sent for protein identification by trypsin digest and shotgun mass spectrometry analysis. It was found to be carbamoyl phosphate synthetase I (CPS1), an enzyme involved in the production of urea in the mitochondria. As the mitochondria and ER are both membrane bound organelles, it is not too surprising about the co-purification of CPS1 and AAT. However, the enrichment of CPS1 in the AATD patient cirrhotic liver was not expected.

6.3.2 Characterisation of *ex vivo* AAT polymers

Due to the impurities in the sample, the extracted polymers were immunocharacterised by western blot analysis with an anti-AAT polyclonal antibody and the polymer-specific 2C1 mAb (Figure 6.3). The SDS-PAGE western blot confirmed that the predominant band in the supernatant was AAT. Furthermore, it suggested that this form of AAT carried immature glycosylation due to its

characteristic faster migration than the plasma monomeric M-AAT and plasma-based heat-induced polymer controls. Lastly, it showed that the liver polymers were non-covalent and in the presence of SDS and reducing agent, dissociate into monomers. In keeping with its polymer specificity, 2C1 did not detect the monomeric forms of AAT. The non-denaturing (native) PAGE showed that all of the AAT in the liver sample was present as polymers. These liver polymers were detected by 2C1, as were the control heat polymers. This supports the notion that the two polymer types have structural similarities and provides additional evidence that 2C1 recognises pathological AAT polymers[76]. Of note is the differing appearance of the heat and liver polymers, where the heat polymers form a laddering pattern whereas the liver polymers are a smear. The liver polymers also appear to be of higher-order than the heat polymers.

A sandwich ELISA was used to determine the concentration of AAT in the liver sample (Figure 6.4). With reference to a heat polymer standard curve and using 3C11 (which recognises total AAT) as the detection antibody, a total amount of 3.2 mg AAT was purified from 3.9 g of liver. When 2C1 was used as the detection antibody, a total amount of 1.3 mg AAT polymer was purified. The reason for the discrepancy could be due to the presence of AAT monomer (which 2C1 does not detect), however monomer was not significantly detected in the western blot. Instead, the difference could represent AAT polymers that are not recognised by 2C1, as has been documented recently[191]. It also may reflect the use of heat-induced polymers as a standard to quantify liver polymers. There would be discrepancies if the polymers differed in their presentation of the 2C1 epitope (for example due to polymer interactions); the 2C1 epitope is situated in the vicinity of helices E and F compared with 3C11, which is near helix A on the opposite side of AAT. Regardless, both these estimates are likely to be an underestimate of the amounts found in unprocessed liver due to sample losses in purification and incomplete polymer release after sonication; when the IB pellets were combined and re-sonicated, remaining AAT in the soluble fraction could still be detected on a denaturing gel.

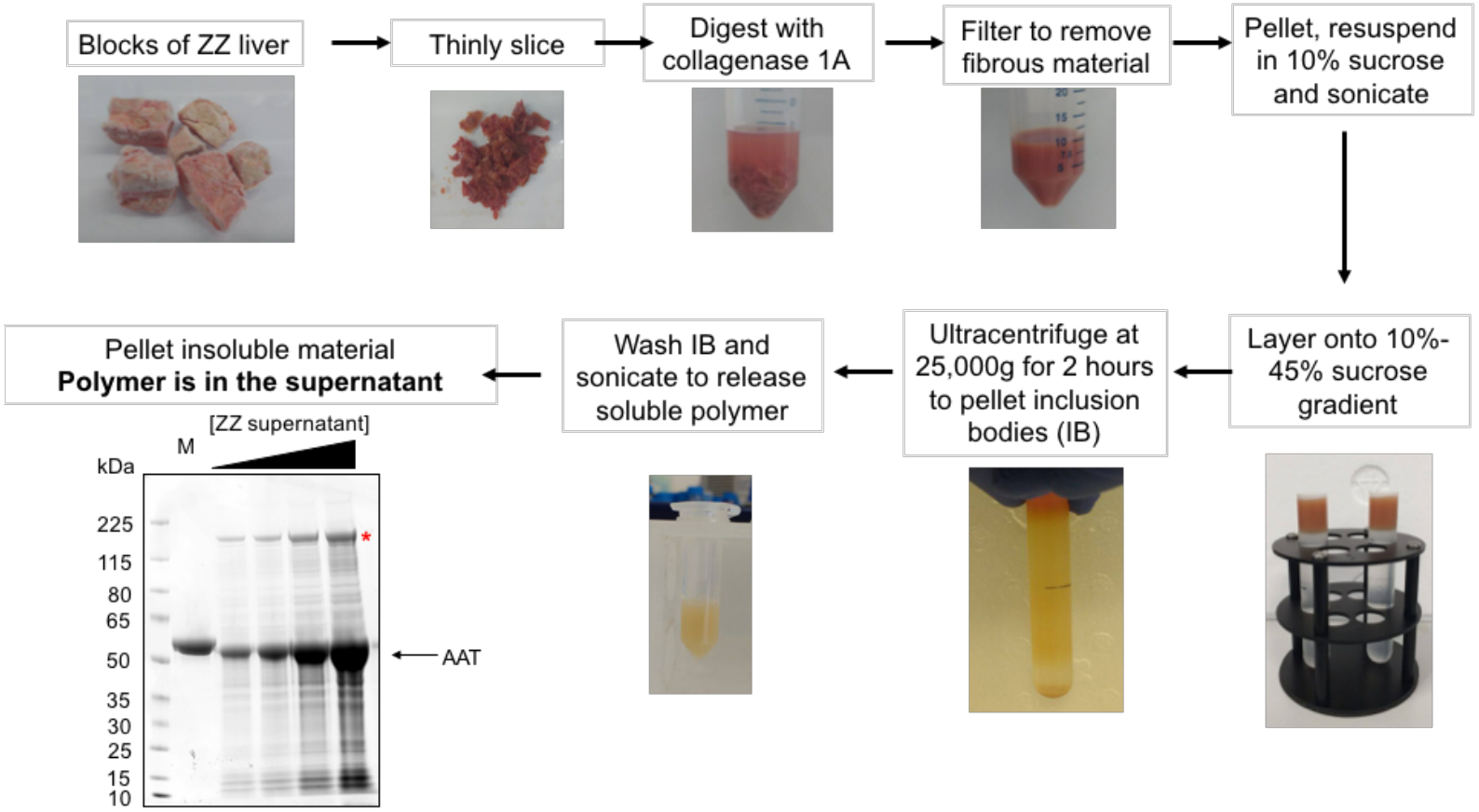


Figure 6.2. Method of polymer extraction from liver samples. *Liver samples were sliced and digested with collagenase. The digest was sonicated after filtering and ultracentrifuged in the presence of a sucrose gradient. The pelleted inclusion bodies were sonicated to release the AAT polymers and the supernatant run on a 4-12% v/v acrylamide SDS-PAGE gel next to a M monomer control. The * denotes CPS1, identified by mass spectrometry.*

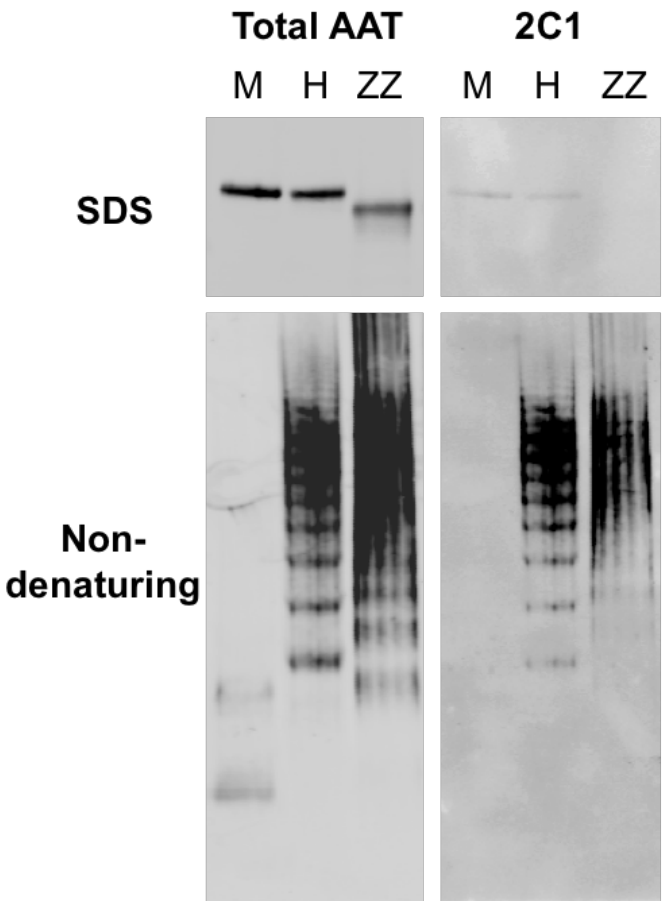


Figure 6.3. Western blot analysis of the liver sample. *The top panels show an SDS reducing gel with 0.4 µg of samples detected by a rabbit total anti-AAT antibody (left) and anti-polymer 2C1 (right). The bottom panels show a non-denaturing gel detected by a rabbit total anti-AAT antibody (left) and 2C1 (right). M corresponds to monomeric M-AAT, H corresponds to heat-induced polymer and ZZ corresponds to polymer isolated from the ZZ homozygote liver. The ZZ sample contains AAT polymer detected by 2C1 and a near-absence of monomer.*

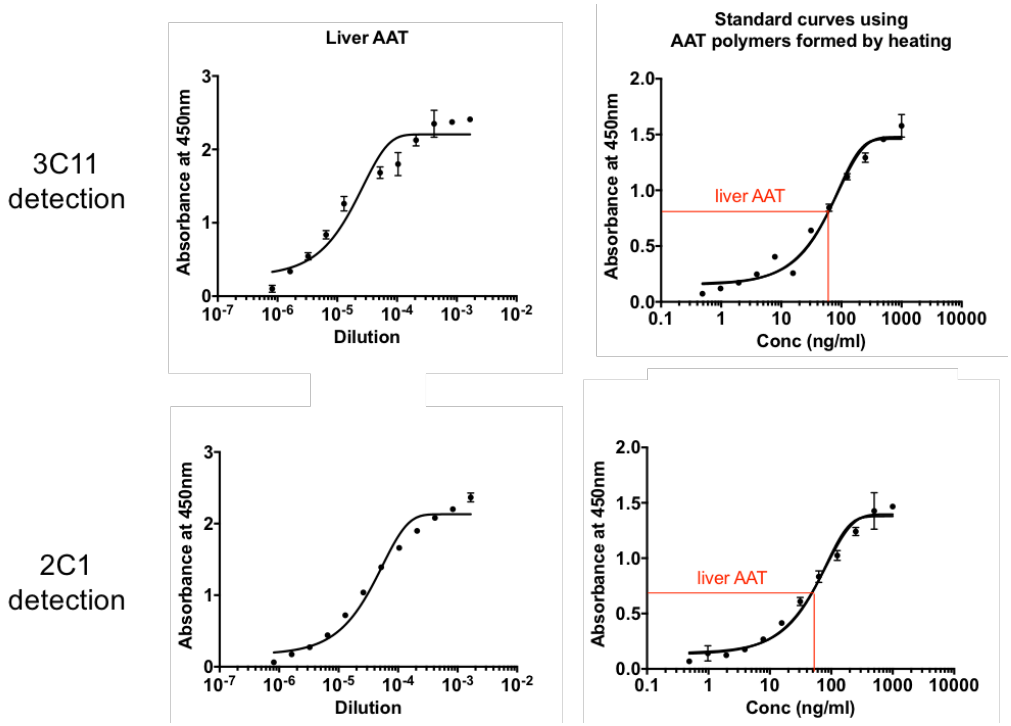


Figure 6.4. Quantification of AAT polymer in the liver sample. *Sandwich ELISAs of ex vivo liver sample (left) and heat polymer standard (right) with 3C11 (anti- total AAT, top) and 2C1 (anti-polymer, bottom) as the detection antibody. Two or three points on the liver AAT graphs were used to interpolate the concentration from the linear region of the standard curves. Each graph summarises three datasets, error bars represent the standard deviation.*

The glycosylation state of the ZZ liver polymers was probed using enzymatic cleavage by the glycosidases, endoglycosidase H (EndoH) and PNGaseF. Glycosylation is the addition of sugars to proteins and occurs within the ER and the Golgi network. AAT is N-linked glycosylated on three residues: Asn46, Asn83 and Asn247[34]. N-linked glycosylation involves the attachment of a large sugar precursor comprised of N-acetylglucosamine (GlcNAc), mannose and glucose to Asn residues in the ER. These sugars are trimmed and added to within the Golgi network to diversify the glycan structures of secreted proteins. EndoH cleaves the glycosidic bond between GlcNAc sugars proximal to the Asn but is very specific for mannose-rich oligosaccharides. PNGaseF on the other hand cleaves the glycosidic bond between GlcNAc and Asn and can cleave both high mannose and complex glycans. Incubation of samples with these two enzymes can therefore be informative about

the type of glycan attached to proteins. ZZ-AAT liver polymers and plasma-derived M-AAT were incubated with the glycosidases and resolved on a denaturing gel (Figure 6.5). ZZ-AAT liver polymers were cleaved by both PNGaseF and EndoH to a shorter deglycosylated AAT whereas plasma monomer was only cleaved by PNGaseF. This indicates that the plasma protein has been processed to the complex glycan form by passage through the secretory system whereas ZZ-AAT liver polymers have high mannose glycans, as would be expected from proteins that have been retained in the ER.

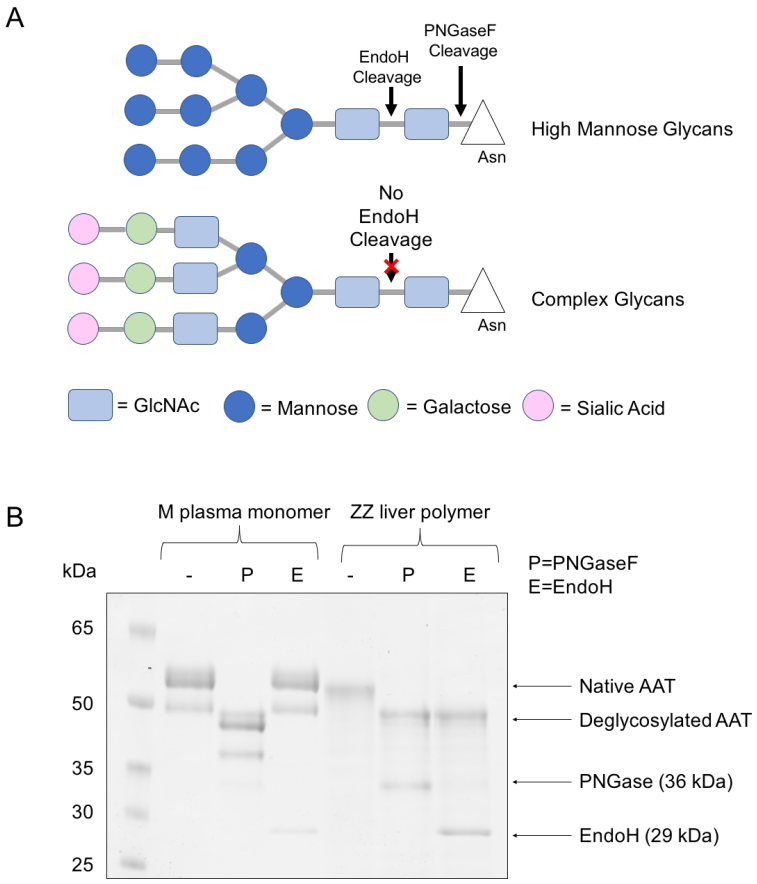


Figure 6.5. Deglycosylation of ZZ-AAT liver polymers. *A) A schematic adapted from [283], illustrating the different glycan structures and their cleavage by EndoH and PNGaseF. B) A 4-12% w/v acrylamide denaturing gel showing cleavage of plasma-derived M-AAT and ZZ AAT liver polymer by P (PNGaseF) or E (EndoH). Each lane contains 2 µg of AAT.*

6.3.3 NS-EM of liver polymers with and without 4B12 Fab

The polymers isolated from the liver explant tissue following sonication of IBs were subjected to dialysis using a membrane with a 300 kDa cut-off and imaged using NS-EM with uranyl acetate. In the micrographs, the polymers appeared long and flexible with individual subunits well-delineated (Figure 6.6). This is similar to the 'beads-on-a-string' appearance previously observed for polymers isolated from the plasma of a patient with the Siiyama AAT mutation[81]. Linear polymers were abundant with occasional circular polymers also present in the liver sample, as noted before. The liver polymers were also similar in appearance to the heat-polymers imaged by NS-EM (Figure 5.3). Polymers were the predominant assembly in the liver sample however there were other irregular stained entities. These are likely to be protein contaminants or ER membrane debris released from the sonication of the IBs.

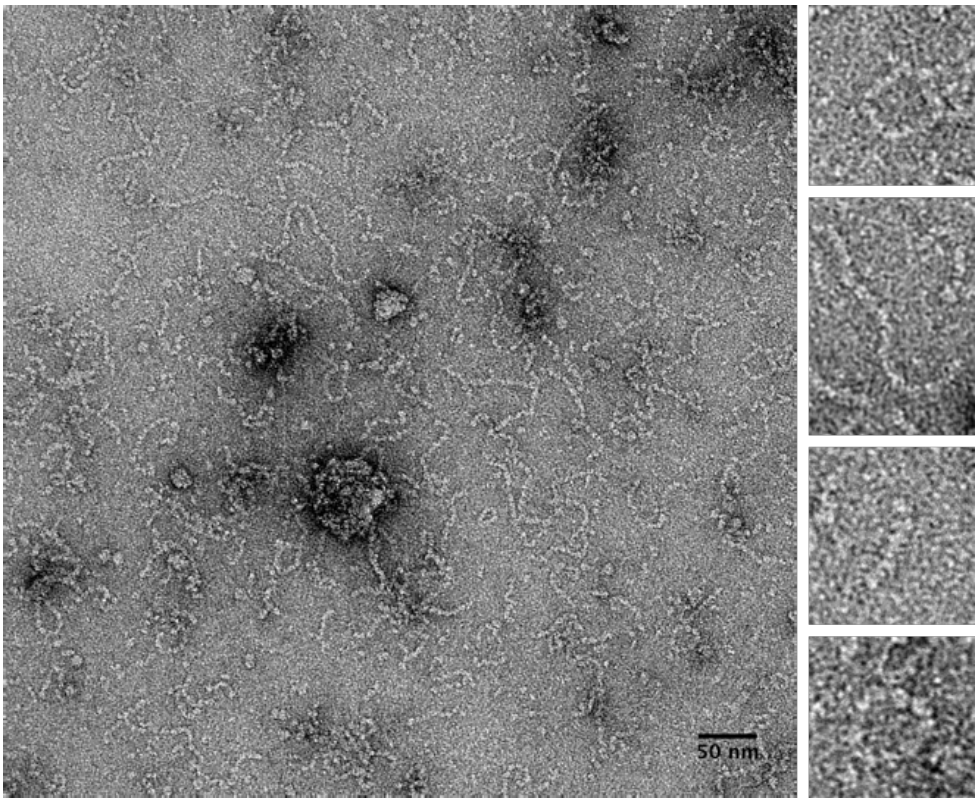


Figure 6.6. A negative stain micrograph of ex vivo liver polymers. *Sample was diluted to a nominal concentration of 0.06 mg/ml in Tris-based buffer, stained with 2% w/v uranyl*

acetate and visualised using a Tecnai T12 microscope at 67,000X magnification. Micrograph taken at 1µm defocus. Scale bar, 50nm.

As for the heat polymers, the *ex vivo* polymers were bound to 4B12 Fab and imaged by NS-EM (Figure 6.7). The Fabs added visible protrusions on the polymer chains providing extra confidence that these were AAT polymers and they were the predominant species present in the liver sample. Most of the 4B12 Fabs could be seen on one side the polymer but there were also polymers with Fabs on different sides. This is similar to what was observed for the heat polymers labelled with 4B12 Fab (Figure 5.5).

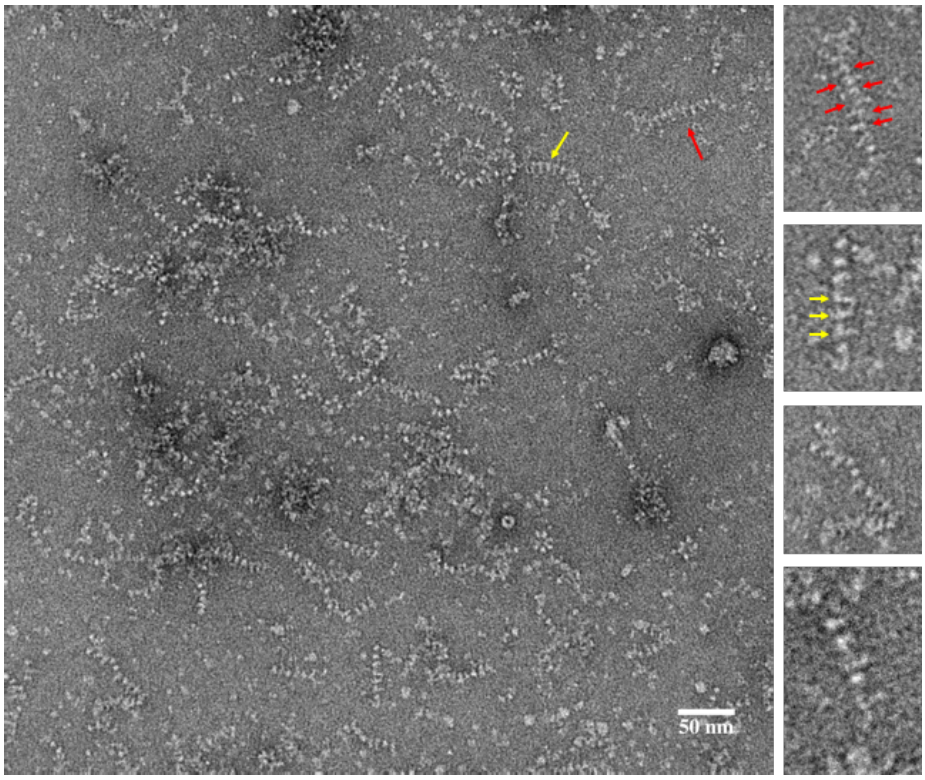


Figure 6.7. A negative stain micrograph of ex vivo liver polymers labelled with 4B12 Fab. *Sample was stained with 2% w/v uranyl acetate and visualised using a Tecnai T12 microscope at 67,000X nominal magnification. Micrograph taken at 1µm defocus. Yellow and red arrows highlight where Fab fragments are lined up on the same or opposite sides of the polymer, respectively. Scale bar, 50nm.*

6.3.4 Approaches to reduce polymer flexibility

Polymer flexibility is likely to be a confounding factor in obtaining high-resolution dimer reconstructions, by reducing the number of usable particles and increasing mis-classification. NS-EM was used to assess whether the addition of different antibody fragments could reduce polymer flexibility. Polymer flexibility had been shown to be challenging in the image processing of the NS-EM reconstructions[250]. Limiting this flexibility would reduce sample heterogeneity and help achieve higher resolution in cryo-EM. Different approaches were trialled using heat-induced polymers as the target sample to spare precious liver polymer samples; any promising conditions could then be attempted with the liver polymers.

The first approach undertaken was to label polymers with Fabs from different antibodies that would have different epitopes and could cause different structural effects on the polymer. Polymers were bound with excess 9C5 Fab, 2C1 Fab, 2E4 Fab or 1A3 Fab. The antibody 1A3 was made in the hybridoma screen discussed in Chapter 4. It was shown to have a preference for polymers by ELISA, similar to 2C1. The antibody 2E4 has been identified in a previous hybridoma screen with a preference of Z-AAT over M-AAT. None of the Fabs tested made a noticeable difference to the flexibility of polymers. Fabs can be seen on alternative sides of the polymer chain and polymer curvature is still apparent. A micrograph of polymers labelled with 1A3 Fab is shown as an example in Figure 6.8.

Next, it was postulated that binding polymers with two Fab molecules (4B12 and 9C5) could increase the linearity of polymers with respect to a single-Fab preparation by 9C5 compounding steric effects. Both 4B12 and 9C5 are known to be able to bind to the AAT subunit at the same time through ELISA competition assays[116]: they do not sterically block each other from binding to AAT and their epitopes must be sufficiently far apart. Heat polymers were labelled with 4B12 and 9C5 Fabs simultaneously and imaged by NS-EM. The polymers were easily visible with Fab-like protrusions away from the main chain. The binding of two Fabs made the polymers more complex in appearance and it was difficult to interpret the relative positioning of the binding sites of the two Fabs on the AAT subunit as the 9C5 epitope was unknown. The binding of the second Fab also did not appear to reduce

the flexibility or increase the linearity of the polymers as tight turns and circular polymers can still be seen.

The last approaches tested were the labelling of polymers with 4B12 mAb or 4B12 Fab_2 fragments. Fab_2 fragments – lacking the Fc component but retaining the bivalency of the parent antibody - are produced by a similar method to the Fab, but mAb digestion with ficin is for longer (24 hours rather than 4 hours) and at a lower cysteine concentration (4mM rather than 25mM). 4B12 Fab_2 production was confirmed by a denaturing gel, where Fab_2 runs around 100 kDa (not shown). Both antibody variations have two binding sites for AAT, so in theory they could bind to two neighbouring AAT molecules within the polymer chain and could cause some rigidity. They were incubated sub-stoichiometrically with polymers (hence the presence of unbound polymers in the micrographs in Figure 6.8) to limit inter-chain binding of the antibody. For both samples, there appeared to be greater clumping of polymers indicating that the two AAT binding sites per 4B12 molecule were binding between chains. Therefore, after all these tests, the 4B12 Fab was retained as the polymer labelling tool for the cryo-EM studies.

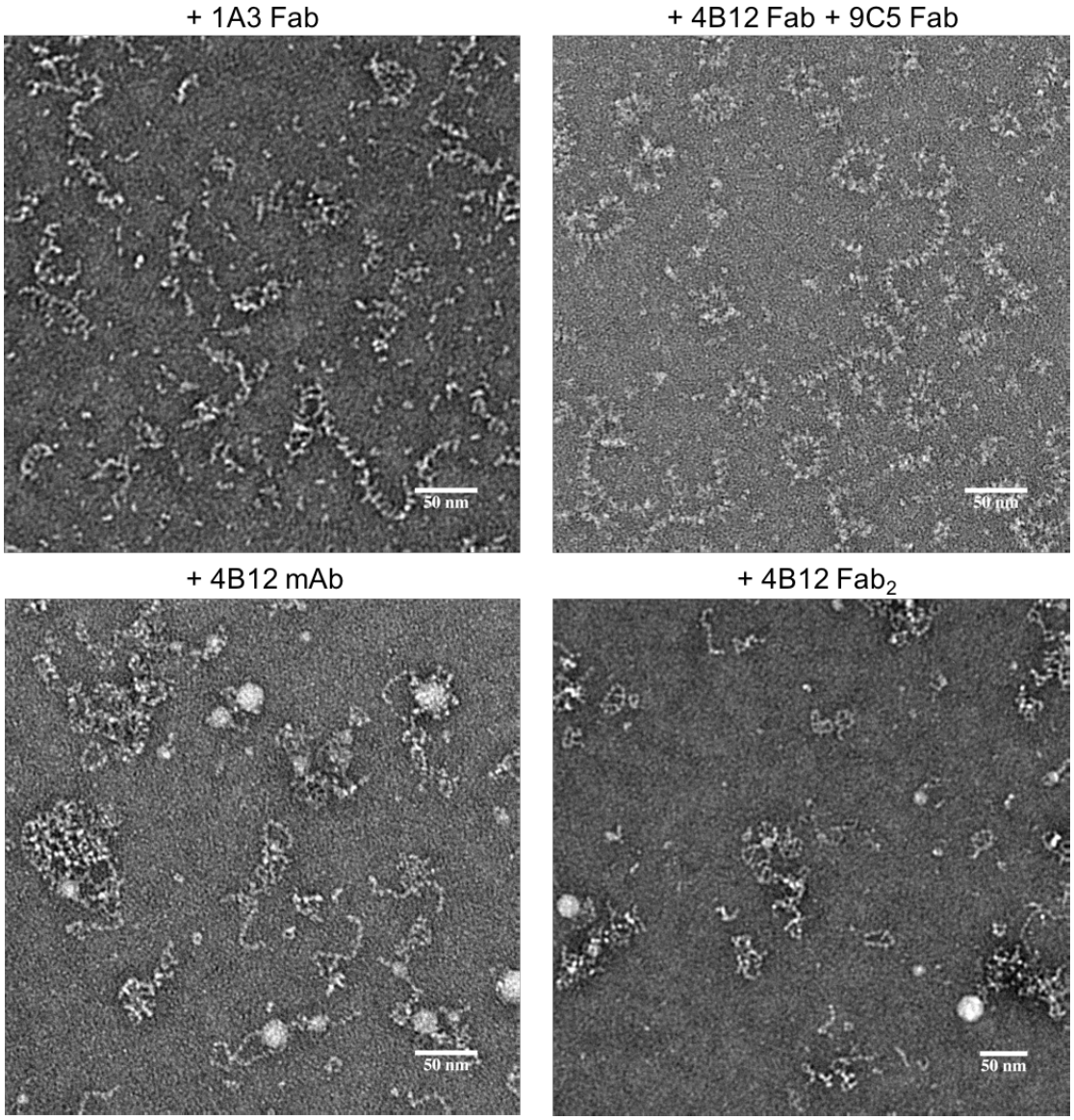


Figure 6.8. Negative stain micrographs of heat-induced polymers labelled with different antibody fragments. *Samples were incubated with either 3 times excess of 1A3 Fab or 4B12 Fab/9C5 or at a 0.5 molar ratio of 4B12 mAb or 4B12 Fab$_2$. Samples were stained with 2% uranyl acetate w/v and visualised using a Tecnai T10 or T12 microscope at 44,000X or 67,000X nominal magnification respectively. Scale bars, 50nm.*

6.3.5 Initial cryo-EM studies

Initial cryo-EM studies were once again carried out using heat polymers labelled with 4B12 Fab to spare the limited liver polymer sample. Sample preparation for cryo-EM is technically more difficult than NS-EM as grids must be quickly plunged into liquid ethane and maintained at liquid nitrogen temperatures for imaging. Vitrification of grids was achieved using a Vitrobot (FEI). There are multiple parameters (grid type, sample concentration, sample incubation time, blot time, blot force etc.) that should be optimised for a sample to achieve the right level of ice thickness; if the ice is too thick the signal will be weak, while thin ice results in particles that tend to lie along the plane of the ice surface.

There are multiple support grids available that may be more or less suitable depending on sample characteristics. The heat polymer/4B12 sample was initially frozen on Lacey carbon grids. Lacey grids have a mesh of carbon film on a copper grid, with different size holes free of carbon in which the sample is imaged. Grids were imaged using a Tecnai 200 kV FEG microscope. Areas of carbon were included in the micrograph images to aid calculation of the CTF.

The sample was first frozen after a 30 second sample pre-incubation, a 5 second blot time and 5 blot force on the Vitrobot. The micrographs showed that the sample was present within the holes of the carbon mesh and so this type of grid was suitable for future work. The contrast by cryo-EM is clearly much lower than NS-EM images however the polymers represent distinctive structures whose path could still be deciphered. Suitable ice thickness for imaging was present, however most of the grid was dry with no sample. To increase the number of carbon holes within the grid that could be used for data collection, the blot time was reduced to 4 seconds. The sample concentration was also reduced three times. The reduction in blotting time increased the number of holes with thin ice and the lowered concentration enabled clearer depiction of the polymer chains. The polymer spread was heterogeneous, as it appeared to cluster near the carbon but could also be seen in the holes. Smaller holes also seemed to have a higher concentration of sample. Images taken at different defoci are shown in Figure 6.9. Choosing the appropriate range of defocus values to acquire images is a balance between obtaining a high enough defocus to

see and pick particles and a low enough defocus to contain high frequency and therefore high-resolution information. At the highest defocus (3.5 μm) the polymer chains are easily visible with a high contrast against the background of ice (zoom, Figure 6.9). As images are taken closer to focus, it gets harder to decipher the polymer chains. From these images, the lowest defocus to confidently pick polymer chains for image analysis is around 1 μm.

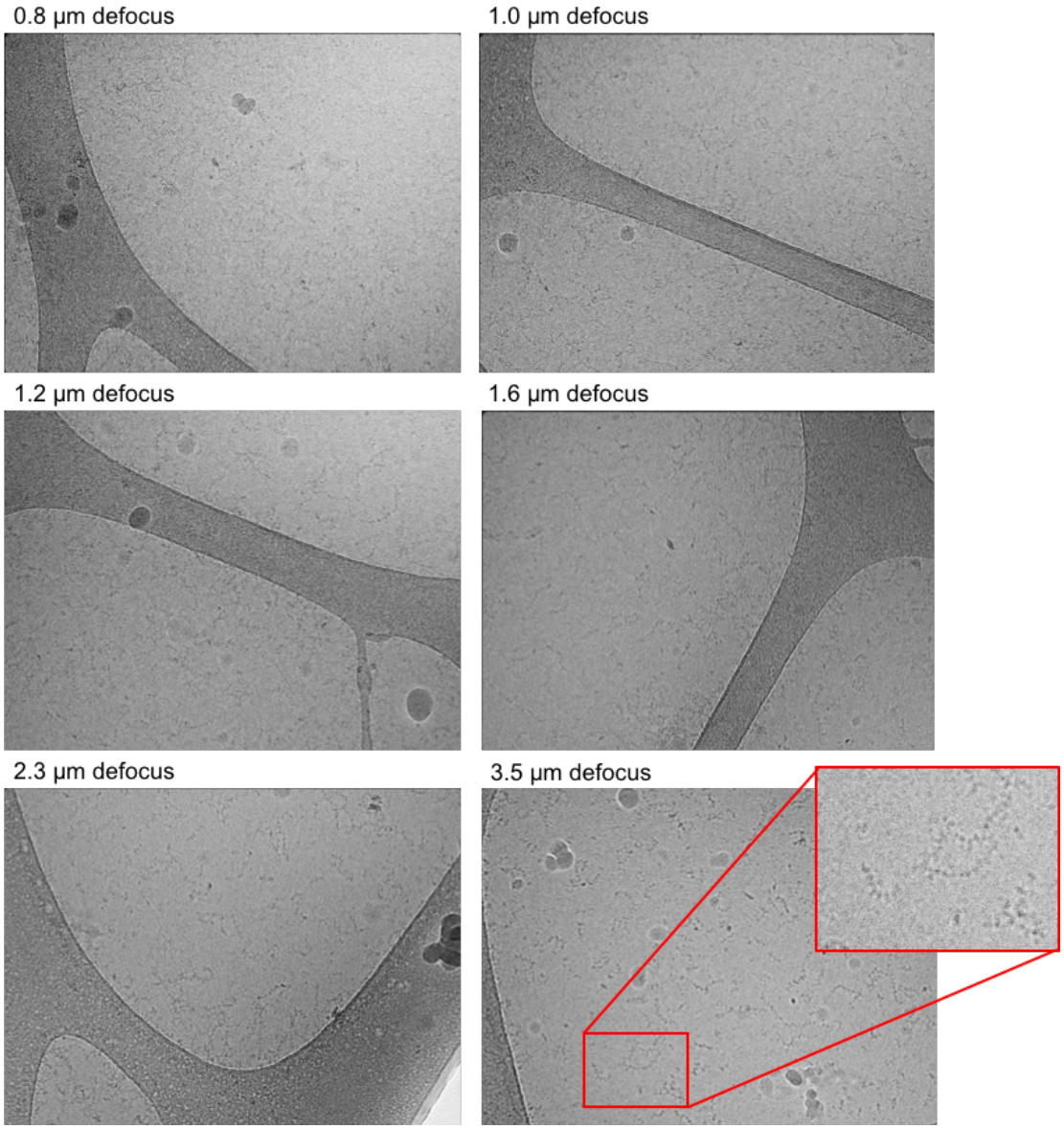


Figure 6.9. Cryo-EM micrographs of heat polymer with 4B12 Fab bound. *Micrographs were taken at a nominal magnification of 41,500X at different defoci on the Tecnai 200 kV microscope. Subframes from 2 to 18 were aligned using EMAN2*[251] *and IMOD*[252] *and corrected for the CTF using CTFFIND3*[253].

As each sample behaves differently when freezing grids, the best freezing condition for the heat polymer/Fab sample could act as a guide, but not necessarily be the best condition for freezing of the liver polymer/Fab sample. Therefore, the freezing protocol of the liver sample was then optimised using Lacey grids. After systematically changing multiple freezing parameters, the ice thickness of the grids and the sample concentration remained highly variable. This irregularity combined with the appearance of non-ice crystals on a low magnification view of the grids (Figure 6.10) suggested that further sample purification might be required. The polymers had only been released from the sonication of IBs with no further purification and non-protein components would not be visible by PAGE. Therefore, it is likely that various cellular debris including lipids were present in the sample being applied to grids, and it may be that these were the source of the crystalline material.

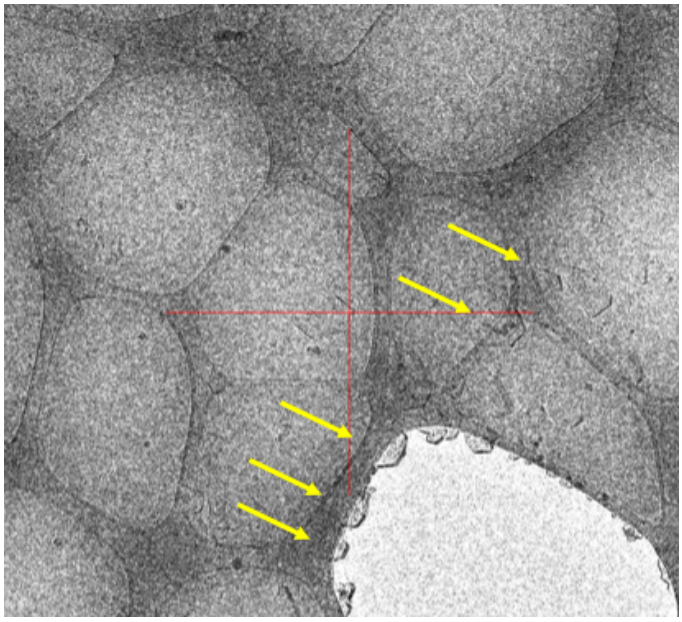


Figure 6.10. Unidentified crystals on Lacey grids of the ex vivo polymer labelled with 4B12 Fab. *Shown is an image from view mode on serial EM, where some crystals are highlighted by arrows.*

6.3.6 Extra purification of *ex vivo* polymers

After the initial cryo-EM studies, it became apparent that to have the best chances of obtaining high-resolution data, the purity of the sample needed to be improved.

Previous attempts of purification using a 1ml Alpha$_1$-Antitrypsin Select column and a custom anti-AAT rabbit polyclonal antibody column, which were both operated by gravity-flow, proved unsuccessful (not shown), as a high proportion of material was lost or did not bind to the column. This was particularly true for the rabbit antibody column where it seemed that the process of conjugation had compromised antibody affinity.

As only small amounts of liver polymer were available, it was difficult to trial many different purification approaches. However, when a high-specification chromatography system became available to use, it was possible to try the Q-sepharose column for purification of the liver polymer sample. This column is used routinely as the second step of AAT purifications and so the behaviour of AAT with the column was well-known. Small amounts of heat-induced polymer were first run to ensure detection of eluted material would be possible; an elution peak could be observed on loading only 50µg of purified protein. Therefore, based on the concentrations measured by ELISA, 150µg of liver polymer sample was loaded onto the Q-column and eluted using a 0-1M NaCl gradient. The absorbance trace showed a reasonably 'messy' elution, reflecting the high number of impurities present within the sample. Denaturing gels were run of the elution peaks: the peak containing AAT is highlighted in Figure 6.11A. Liver AAT eluted at 300 mM in the salt gradient compared to 340 mM for heat-induced polymers made from plasma monomeric AAT. The peak for the polymer was broad but only fractions in the defined peak were pooled for labelling by 4B12 Fab. The first peak in the elution contained a low molecular weight contaminant (gel not shown) and the third large peak was not visible on a gel, which might be the result of non-protein contaminants such as lipids that absorbed in the UV range. The purity of the resulting AAT polymer sample was much higher and this was apparent from NS-EM micrographs of the purified sample, where very little non-AAT debris could be seen (Figure 6.11B).

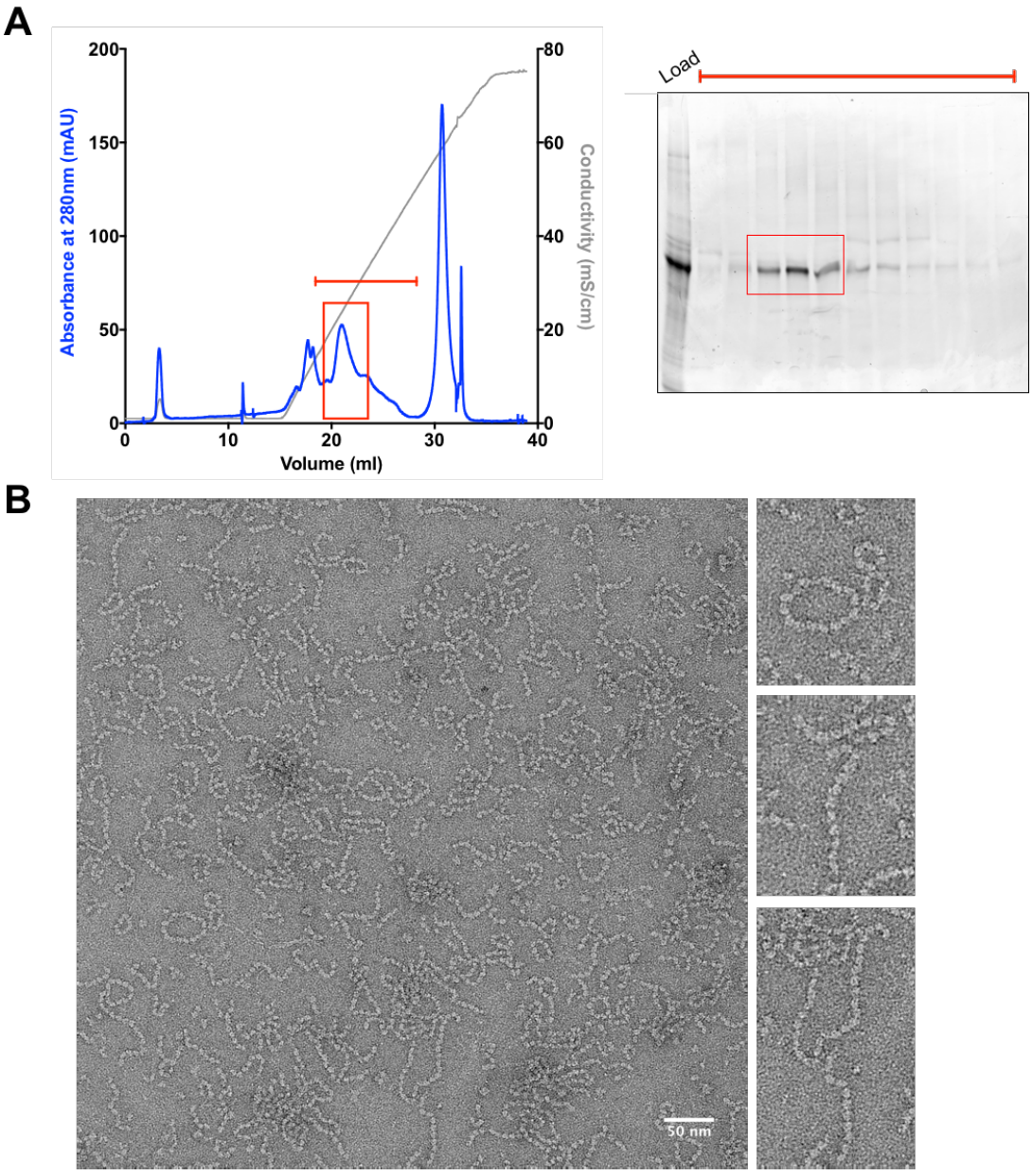


Figure 6.11. Extra purification of ex vivo polymers. *The elution profile (left) of the ex vivo sample from a Q-sepharose column. The trace in blue is the UV absorbance at 280nm, the conductivity (a proxy for salt concentration) is in grey. The peak highlighted by a red box and line was resolved by a 4-12% acrylamide denaturing gel (right) by loading 10 µl samples; these fractions corresponded to AAT. The first lane in the gel labelled 'load' was the impure*

sample before purification by the Q-sepharose column. B) The sample was visualised by NS-EM with 2% w/v UA staining on a Tecnai T12 microscope at 67,000X nominal magnification.

The previous approach to purifying labelled polymers consisted of dialysis using a membrane with a 300kDa molecular weight cut-off, following incubation with an excess of 4B12 Fab to remove unbound Fab. However, small particles could still be seen by NS-EM and so the efficiency of this approach was questionable (Figure 5.5). After observing that the AAT polymers from the liver could be bound by Q-column, it was hypothesised that this step could be used a second time in place of the dialysis step. This would require a lack of binding by the 4B12 Fab under the buffer conditions used or binding with a different strength to AAT (such that the Fab would elute at a different point in the salt gradient). It was also important to understand the effect of the 4B12 Fab on the concentration of salt at which the complex eluted.

As a test run, small amounts of monomeric M-AAT from plasma and 4B12 Fab were bound and eluted from the Q-column on their own, followed by the monomer AAT/4B12 Fab pre-formed complex. The elution profiles are overlaid in Figure 6.12A, alongside corresponding denaturing gels to identify which protein was present within each fraction. The elution of the M-AAT sample showed a small peak followed by a much larger second peak. When corresponding fractions were run on a gel, no protein band was seen in the first peak and pure AAT monomer (with a little dimer) was in the second peak. The maximum of the peak for AAT was at a salt concentration of 260mM. The 4B12 Fab bound to the Q-column as well, so was negatively charged at pH 8. It was eluted from the column in a shouldered peak at a maximum salt concentration of 120mM. The 4B12 Fab therefore bound less strongly than AAT, which suggested that excess Fab could be purified away after incubation with AAT polymer. The pre-formed complex showed this to be true, where excess 4B12 Fab was apparent from the denaturing gel in the first peak and both AAT/Fab (the complex) were present in the second larger peak. This also demonstrated that 4B12 Fab binding to AAT did not alter the charge of AAT significantly as it was eluted at a similar salt concentration of 270mM.

After establishing that the Q-column was a viable option to purify labelled polymers, the liver AAT polymers were incubated with a three-fold 4B12 Fab molar excess and

then bound to the Q-column. The elution absorbance profile showed two clear peaks; when these were analysed by SDS-PAGE, the first was clearly free 4B12 Fab and the second was AAT labelled with 4B12 Fab in a 1:1 stoichiometry (Figure 6.12B). As there was free Fab, there was confidence that the AAT polymer was fully labelled and the purity of the sample was much higher. The improved sample was reassessed by NS-EM prior to commencing cryo-EM studies again (Figure 6.12C).

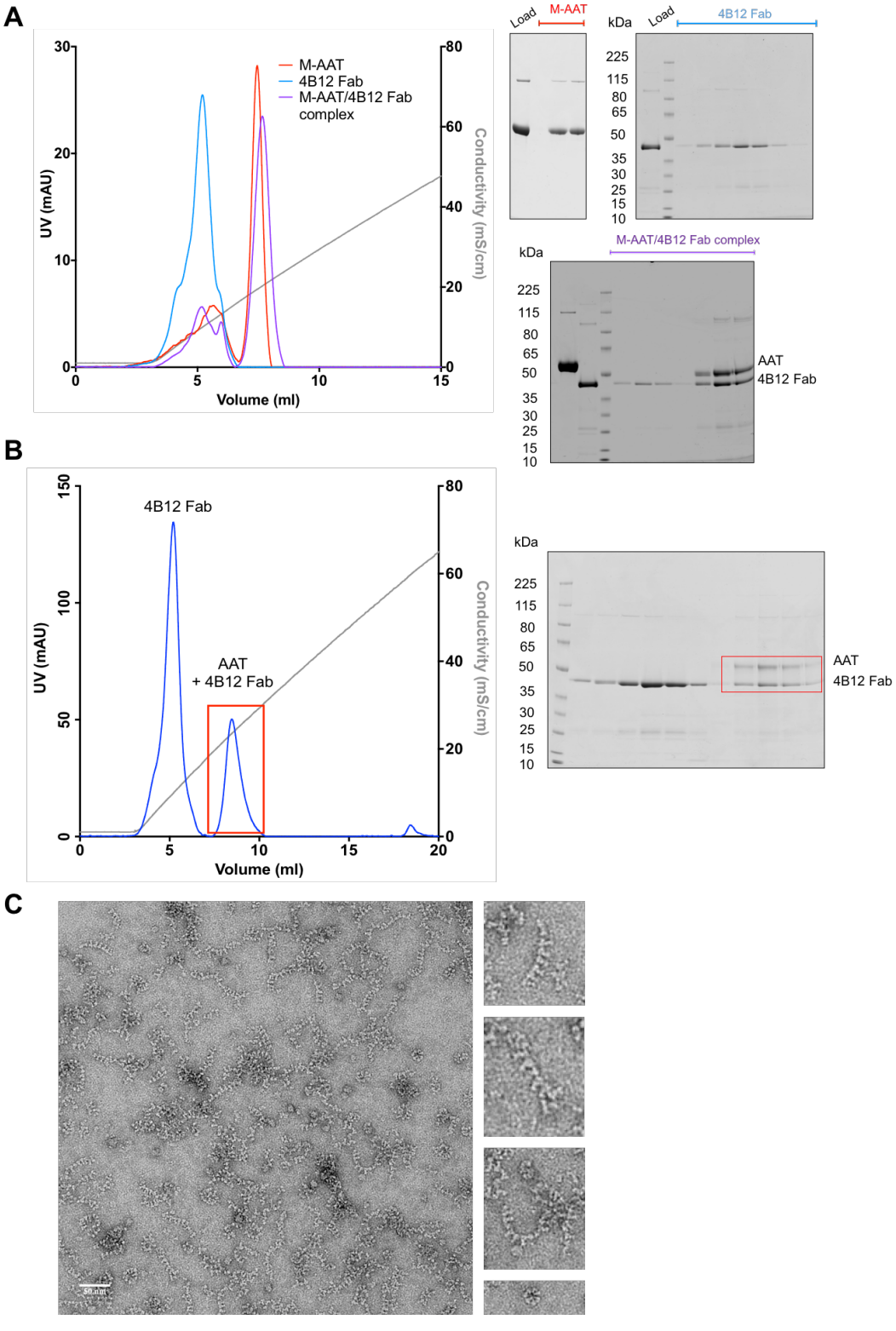

A
M-AAT
4B12 Fab
M-AAT/4B12 Fab complex
UV (mAU)
Conductivity (mS/cm)
Volume (ml)
Load
M-AAT
kDa
Load
4B12 Fab
M-AAT/4B12 Fab complex
AAT
4B12 Fab
B
4B12 Fab
AAT
+ 4B12 Fab
UV (mAU)
Conductivity (mS/cm)
Volume (ml)
kDa
AAT
4B12 Fab
C

Figure 6.12. Labelling of ex vivo polymers with 4B12 Fab. *A) The overlaid elution profiles (left) of monomeric M-AAT (red), 4B12 Fab (blue) and M-AAT/4B12 Fab complex (purple) from a Q-sepharose column. Corresponding 4-12% acrylamide denaturing gels loaded with 20 µl of sample. B) The elution profile of ex vivo polymer labelled with excess 4B12 Fab from a Q-sepharose column. The trace in blue is the UV absorbance, the conductivity measure is in grey and the peak highlighted by a red box corresponds to AAT bound with 4B12 Fab, shown by the denaturing gel (right). C) The sample was visualised by NS-EM with 2% w/v UA staining on a Tecnai T12 microscope at 67,000X nominal magnification.*

6.3.7 Cryo-EM grid optimisation

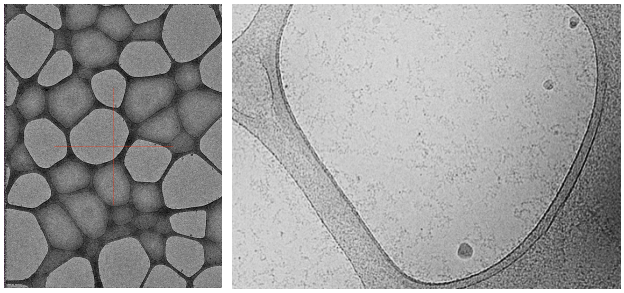
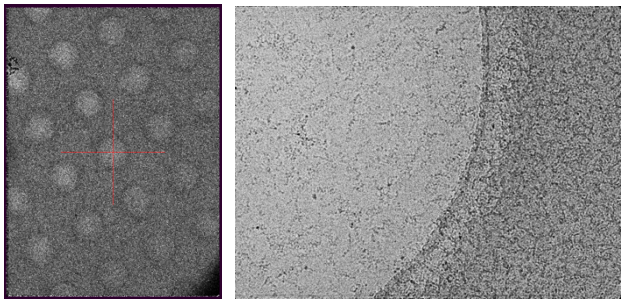
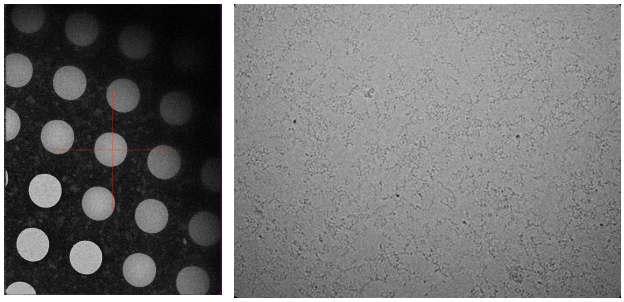
After improving the purity of the sample, optimisation of cryo-EM grids was recommenced. This was approached in a systematic way, optimising the grid type, sample concentration, blot force and blot time in order (Table 6.1). All grids were imaged using the Tecnai 200 kV FEG microscope.

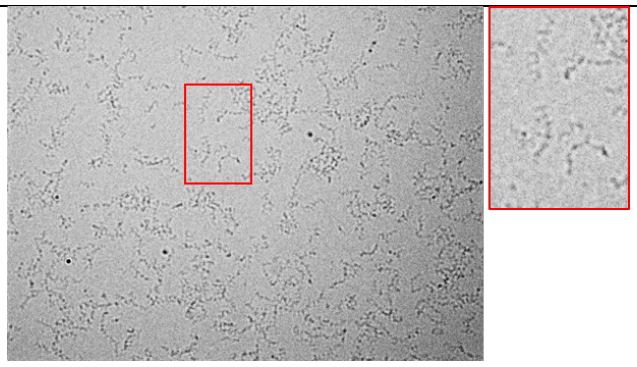
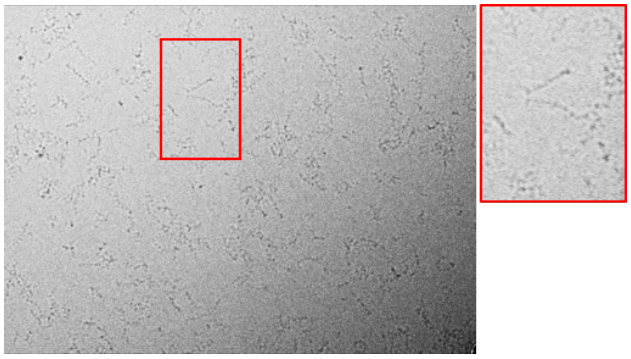
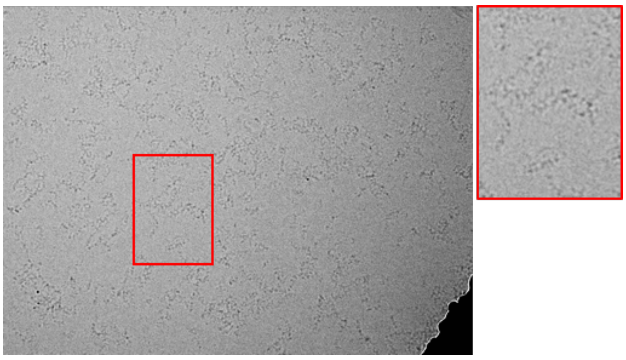
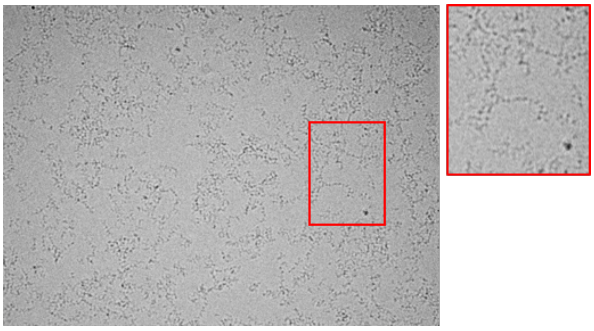
As starting freezing conditions, on recommendation from other cryo-EM users, the purified liver polymer/4B12 sample was applied at a concentration of 0.7 mg/ml with Vitrobot settings of 30 second sample incubation, 5 blot force and 5 sec blot time. The first parameter that was varied was the grid type and three grids were tested. The first type was the Lacey carbon grids that had previously been used in Figure 6.9, which have different sized holes surrounded by carbon. At the conditions used, the larger holes were free from ice but in smaller and medium-sized holes the sample was present across the hole, if not slightly clustered near the carbon edges (Table 6.1). Quantifoil 1.2/1.3 carbon grids were the second grid type tested, which have regular sized holes surrounded by carbon[284]. The polymers were also present within the holes and on the carbon edges. The last grid type was UltrAuFoil 1.2/1.3 gold grids that also have regular holes but no carbon[285]. Again, the polymer chains could clearly be seen within the holes. All grids would therefore be compatible with the sample to further optimise conditions for data collection. However, the advantages and disadvantages of the grid types were considered in deciding on the grid type to proceed with. The Quantifoil grids with regularly spaced holes are much easier for automated data collection and so the Lacey grids with irregular holes were discounted. Gold grids have been demonstrated to be more stable in the electron

beam with 50-fold less movement parallel to the beam than carbon Quantifoil[285]. This improves the quality of micrographs for image processing. However, they do not have any carbon to aid with focusing or CTF correction and so sample concentrations have to be higher to provide enough signal in the Fourier transform for CTF estimation. The gold grids were chosen to maximise the chances of obtaining high-resolution data.

The next parameter that was optimised was the sample concentration. It is desirable to have the sample concentration as high as possible to maximise particle numbers per micrograph in data collection. However, it is also important to minimise the overlap of particles and visualise individual polymer chains for particle picking and image processing. In the testing of grid type, although the sample concentration varied between holes, overall the concentration was too high for particle picking. The concentration was reduced to 0.5 mg/ml. Micrographs showed much better polymer chain delineation with less overlap and the particle number per micrographs was still maximised (Table 6.1). The sample concentration was also high enough to provide adequate signal for CTF correction.

Over half of the ice overage across the grid was thick and therefore the Vitrobot settings of blot force and blot time were next varied to obtain more suitable areas for data collection. Increasing the blot force did reduce ice thickness. The blot force was increased to 10 and 15; 10 blot force was deemed the best as at blot force 15, areas of the grid, particularly in the centre of squares, became dry. Other areas showed that thinner ice was pushing the sample into higher concentration regions. Finally, the ice thickness was fine-tuned by optimising the blot time to 3 seconds. This decreased the number of empty holes that were present in the centre of squares.

<table>
<tr><td colspan="3">Starting conditions: 0.7 mg/ml sample, 30s sample incubation, 5 blot force, 5 sec blot time</td></tr>
<tr><th>Parameter</th><th>Variations</th><th>Outcome</th></tr>
<tr><td rowspan="3">1) Grid type</td><td>a) Lacey grids
  </td><td rowspan="3">c) Gold</td></tr>
<tr><td>b) Quantifoil
  </td></tr>
<tr><td>c) Gold
  </td></tr>
<tr><td rowspan="2">2) Sample concentration</td><td>a) 0.7 mg/ml
(tried above in 1c)</td><td rowspan="2">b) 0.5 mg/ml</td></tr>
<tr><td>b) 0.5 mg/ml</td></tr>
</table>

		
3) Blot force	a) 5 blot force (tried above in 2b)	b) 10 blot force
	b) 10 blot force 	
	c) 15 blot force 	
4) Blot time	a) 3 sec 	a) 3 sec blot

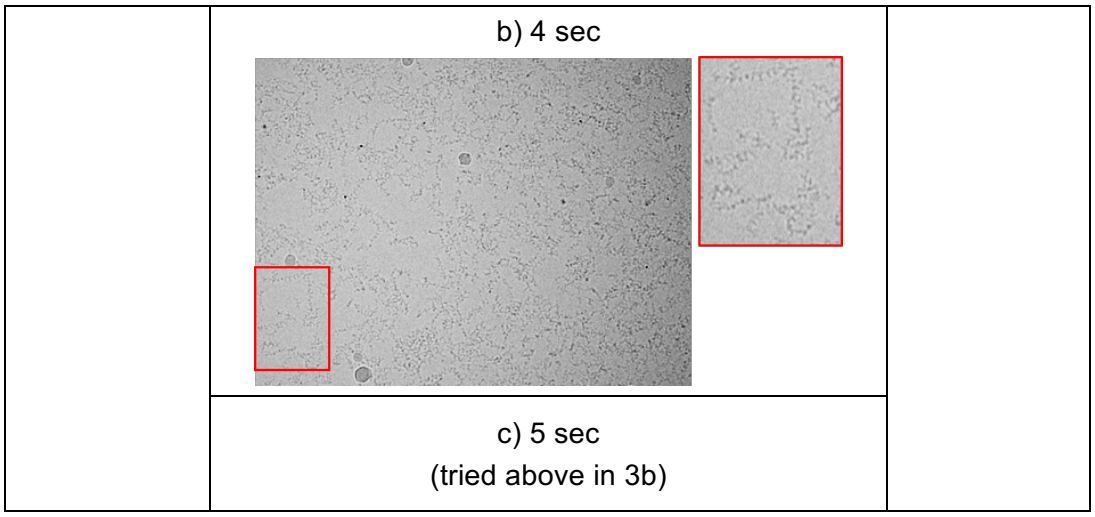


Table 6.1. Summary of cryo-EM grid optimisation for the liver polymer/4B12 Fab sample. *Grid type, sample concentration, blot force and blot time were varied sequentially. Micrographs were taken at a nominal magnification of 29,000X on the Tecnai 200 kV microscope. Subframes 2-17 were aligned EMAN2*[251] *and IMOD*[252] *and corrected for the CTF using CTFFIND3*[253]*.*

6.3.8 Initial cryo-EM dataset collection

Once suitable grid-freezing conditions for the gold grids were established, an initial small dataset was collected on the 300kV Tecnai Polara microscope equipped with a K2 direct detector, and with the help of Dr N. Lukoyanova. An overview of the ice thickness of the squares across the entire grid was captured through a low magnification montage map (LMMM). One example square is shown in Figure 6.13A. Holes towards the edge of the square had thicker ice and there were some thinner-ice or empty squares in the centre. From the LMMM, around 20 squares were selected for collection of a medium magnification montage map (MMMM) to yield more detailed information about ice thickness. An example of even, thin ice is shown in Figure 6.13B and was taken from the centre of a square. The MMMM was used to draw polygons around holes that were suitable for data collection. One micrograph per hole was taken during automated data collection at a magnification that yielded a pixel size of 1.1Å. A representative micrograph (low-pass filtered to aid visualisation) in Figure 6.13C shows that the individual polymer chains can be easily made-out

over the background however, clumping together of chains is also apparent. The concentration of polymer across the grid also varied considerably.

The dataset was processed using Relion v2.1[122]. A total of 60 subframes per micrographs were collected, with a dose of 0.95e^-/Å^2/frame (total dose of 57e^-/Å^2) and these were aligned using dose-weighting in MotionCorr2. The CTF was estimated by CTFFIND4. The match of the Thon rings between the Fourier transform of the micrograph and the theoretical CTF was visually inspected. Those that did not correlate were selected, and the parameters altered to aid re-fitting. If this still did not work, the micrographs were disregarded. An example of a good fit is shown in Figure 6.13D, with Thon rings calculated to extend to 5Å.

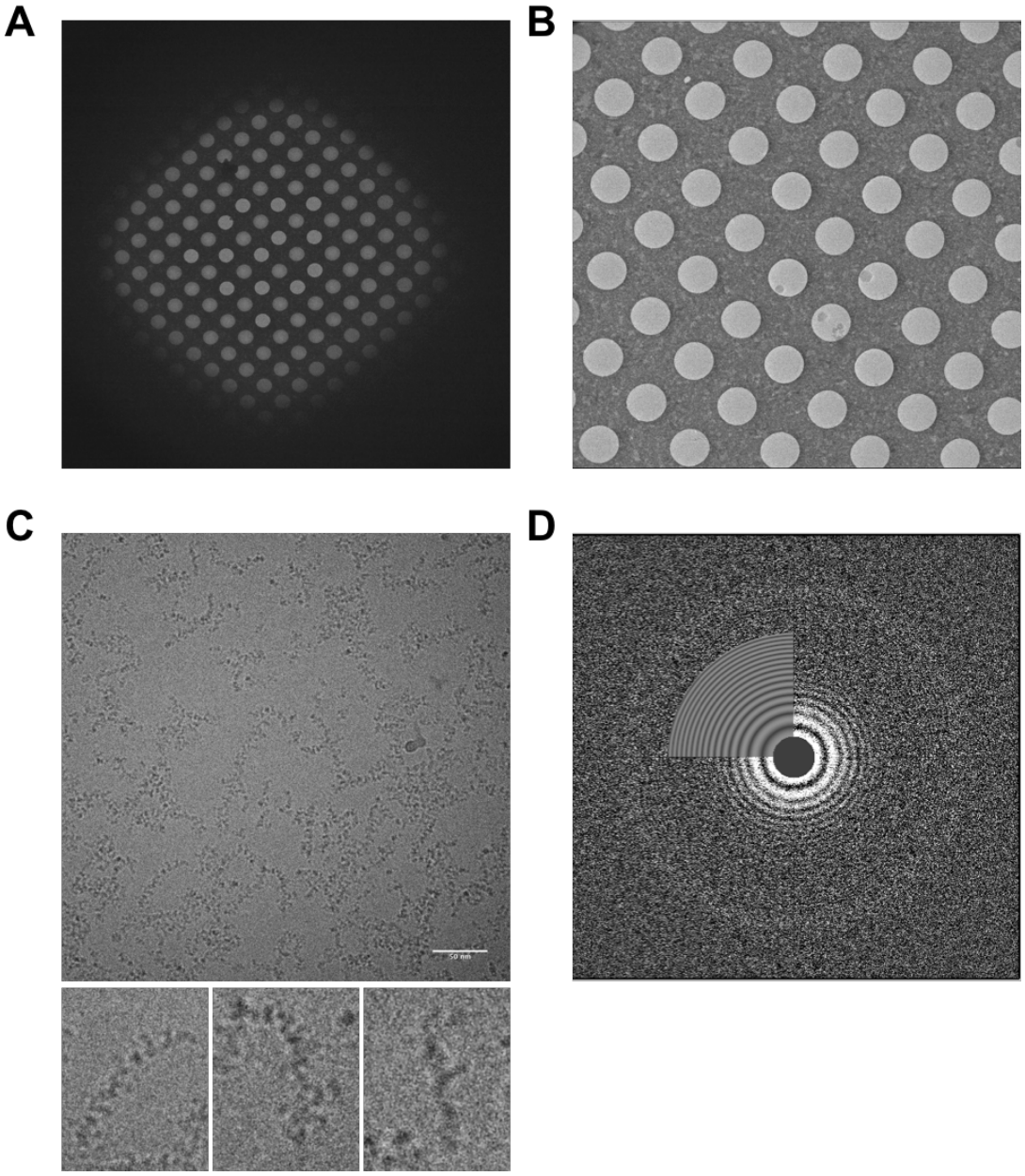


Figure 6.13. Preliminary data collection. *A) The low magnification montage map B) The medium magnification montage map C) A representative micrograph with a 50nm scale bar taken at nominal magnification 200,000X on the FEI Tecnai 300kV Polara microscope equipped with a K2 direct detector. Micrographs were taken over a 12s exposure with a total dose of 57e-/Å^2 and the pixel size was 1.1Å. The 60 subframes were aligned using Motioncor2 within Relion v2.1*[122]*. Example polymers are boxed below. D) An example Fourier transform of a micrograph with the fit to a CTF parameter, calculated using CTFFIND-4.1.*

6.3.9 Processing workflow for the AAT/Fab dimer

As for the NS-EM work, the dimer – the minimal polymeric unit – was selected for reconstruction to provide information on the linkage between subunits in the polymer. Around 1,000 clear dimer particles were manually picked from full-dose micrographs that had been motion- and CTF-corrected. The full-dose micrographs had higher contrast which helped with the manual picking and the coordinates of particles were then applied to lower-dose micrographs ($25e^-/Å^2$) to obtain particles in a 180x180 pixel box with less particle damage. The particles were sorted by reference-free 2D classification into 10 classes (Figure 6.14). The class averages showed two subunits from a chain, each subunit of varying shape depending on the Fab position. Four classes (circled in Figure 6.14) were chosen that represented different views of the dimer and were used as references for autopicking. As the references were generated from the dataset, they would have the same grey intensity scale and would aid the autopicking function within Relion[139]. Initially, parameters for autopicking were set to a relatively high threshold to obtain the best particles in the dataset. These were optimised on a representative subset of four micrographs that varied in defocus and concentration of sample. The ~100,000 particles were sorted by 2D classification and manual subset selection, to yield a final 16,000 particles that formed classes with well-defined dimers (Figure 6.14). The class averages showed the presence of Fab on some AAT subunits and a heterogeneity in the rotation of subunits comparable to that observed with the NS-EM studies. The distance between the centre of masses of subunits was measured, using ImageJ, to be on average 66Å.

The 16,000 particles were used in maximum likelihood projection matching using one of the NS-EM *ex vivo* dimer models as a reference in 3D classification. At the time this was carried out, the Relion NS-EM work had not been completed and therefore a small-angle dimer from the IMAGIC processing was low-pass filtered to 50Å and used as the reference. The particles were divided among four models in an attempt to account for 3D heterogeneity, yet all models only had one adequately formed subunit in the dimer (Round 1, Figure 6.14). Model 3 stood out as having a better-formed subunit: when a model of AAT and 4B12 Fab were docked in to the density in

Chimera (top right of model 3, Figure 6.14) the shape complementarity was high. The region occupied by the 4B12 Fab contained a hole in the density positioned between the variable and constant domains of a Fab. This detail had not been observed in the NS-EM work and represented an unbiased characteristic of the Fab structure, indicating the data were converging on a physically reasonable solution.

This model was used to create a new reference, now constructed from cryo-EM particles, for a second round of 3D classification against those same particles. It was hypothesised that the second subunit was not defined due to the NS-EM dimer failing to adequately represent the angle between the subunits in the cryo-EM data. The reference was created by duplicating and fitting the well-formed subunit on the poorly-formed subunit using Chimera. The reference was filtered to 50Å, effectively removing higher resolution features such as the 'hole' in the Fab density. For most of the generated models, a similar result of only one subunit forming was observed (Round 2, Figure 6.14). However, model 3 again did have density for the second subunit and encouragingly when the AAT/Fab atomic model was docked, the Fab hole was now also apparent for the top subunit. The distance and rotation between subunits were measured to be 65Å and 27°, respectively. Formation of the second subunit suggested this inter-subunit angle was present within the data.

Autopicking had previously been carried out using a small number of references with a high picking threshold. The new cryo-EM model was re-projected on the Euler belt, and the projections used to re-pick the dataset with a lower threshold and therefore maximise the number of particles for processing. Projections that corresponded to end views and hence appeared as a single subunit were excluded from the references to focus on the dimer. Around 250,000 particles were picked and were sorted by 2D classification over several rounds to yield a final 108,000 particles.

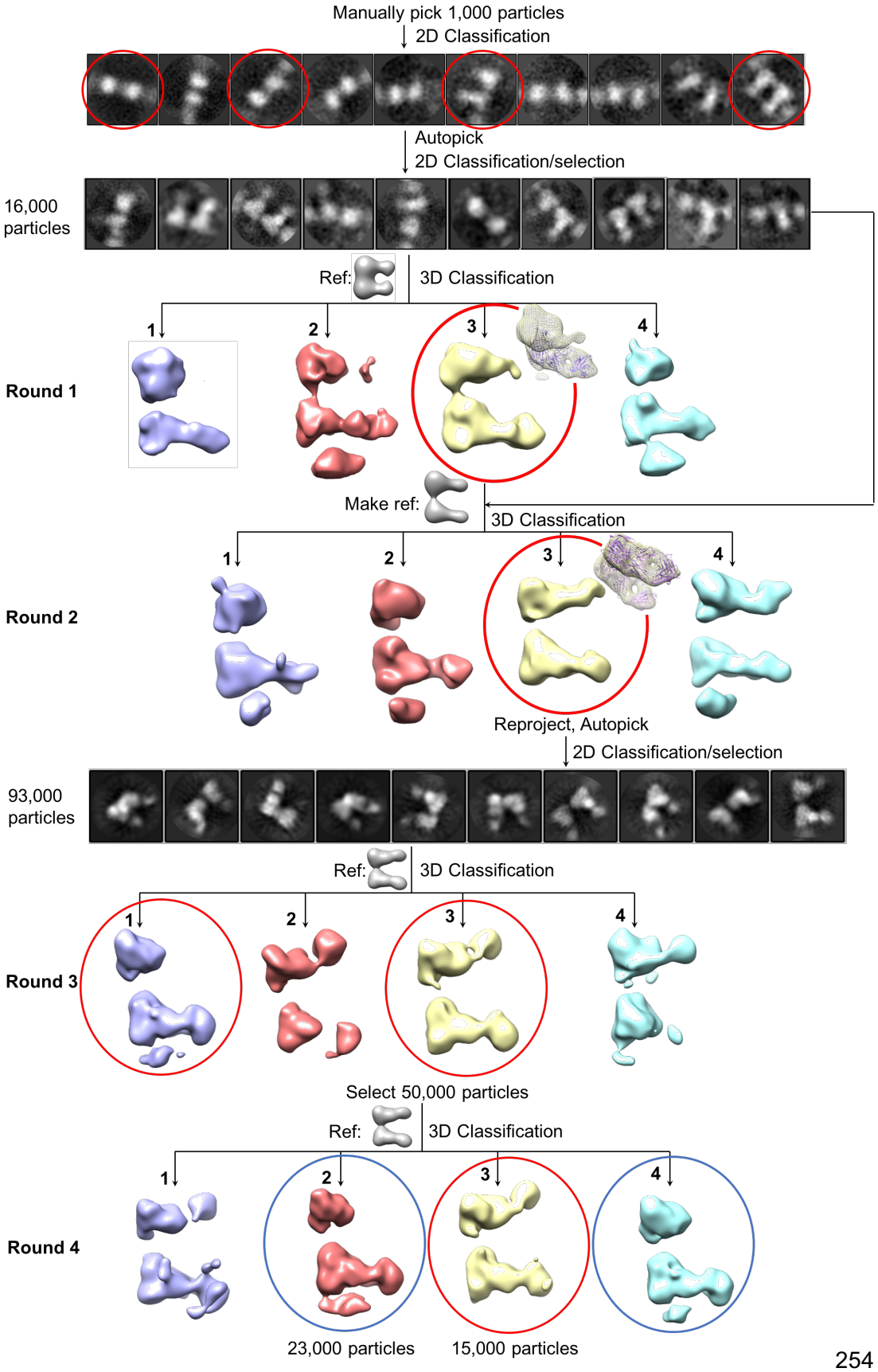

Manually pick 1,000 particles
2D Classification
Autopick
2D Classification/selection
16,000 particles
Ref:
3D Classification
Round 1
1
2
3
4
Make ref:
3D Classification
Round 2
Reproject, Autopick
2D Classification/selection
93,000 particles
Ref:
3D Classification
Round 3
Select 50,000 particles
Ref:
3D Classification
Round 4
23,000 particles
15,000 particles

Figure 6.14. Cryo-EM workflow for image processing of the dimer. *References for autopicking were generated by manually picking 1,000 particles followed by 2D classification and selection. The whole dataset was autopicked for particles and these were sorted by 2D classification to yield a final 16,000 particles. These were subjected to 3D classification using a NS-EM model as a reference. One model was selected based on good subunit definition and used to generate a reference for a second round of 3D classification. Model 3 was selected to be the best and was reprojected around the Euler sphere to produce reprojections for autopicking. After 2D classification and selection, 93,000 particles were subjected to 3D classification. Particles belonging to models 1 and 3 were selected for another round of 3D classification. All reconstructions were thresholded to a volume corresponding to their molecular weight using Chimera (200kDa x 1.2Å^3/Da).*

The 2D class averages represented a wide range of views of the dimer and multiple populations of dimer varying in the inter-subunit rotation (Figure 6.14). Most noticeable, were two class averages that appeared to have a near-180° angle between subunits such that the Fabs were on opposite sides of the AAT chain (Figure 6.15A). This conformation had been observed previously in NS-EM of the heat-induced polymer and *ex vivo* polymer. A simple overlay of two copies of the AAT/Fab atomic model on one of the projections showed excellent shape correlation. On close inspection, it appeared as if more internal details could be made-out in comparison to class averages where the subunits had a smaller rotation between them. The arrows in Figure 6.15A point to three circular areas of higher electron density that could, based on the superimposition, correspond to helices A, G and I in AAT. This conformation might rigidify the dimer allowing higher-resolution analysis. Furthermore, the classes highlight that the polymerisation model must have enough flexibility to accommodate a 180° flipping of the AAT chain. A quick assessment was carried out by perturbing the three proposed models using PyRosetta[164] to yield a dimer with a +180 or -180° rotation between subunits (Figure 6.15B). Purely based on visual inspection, the loop-β sheet A model did not appear to accommodate such an angle in a conformation that would yield the projection view apparent in the class averages. The linkage meant that one subunit of the pair was rotated to a position that would represent a top view in a projection. The β-hairpin and C-terminal models appeared equally accommodating.

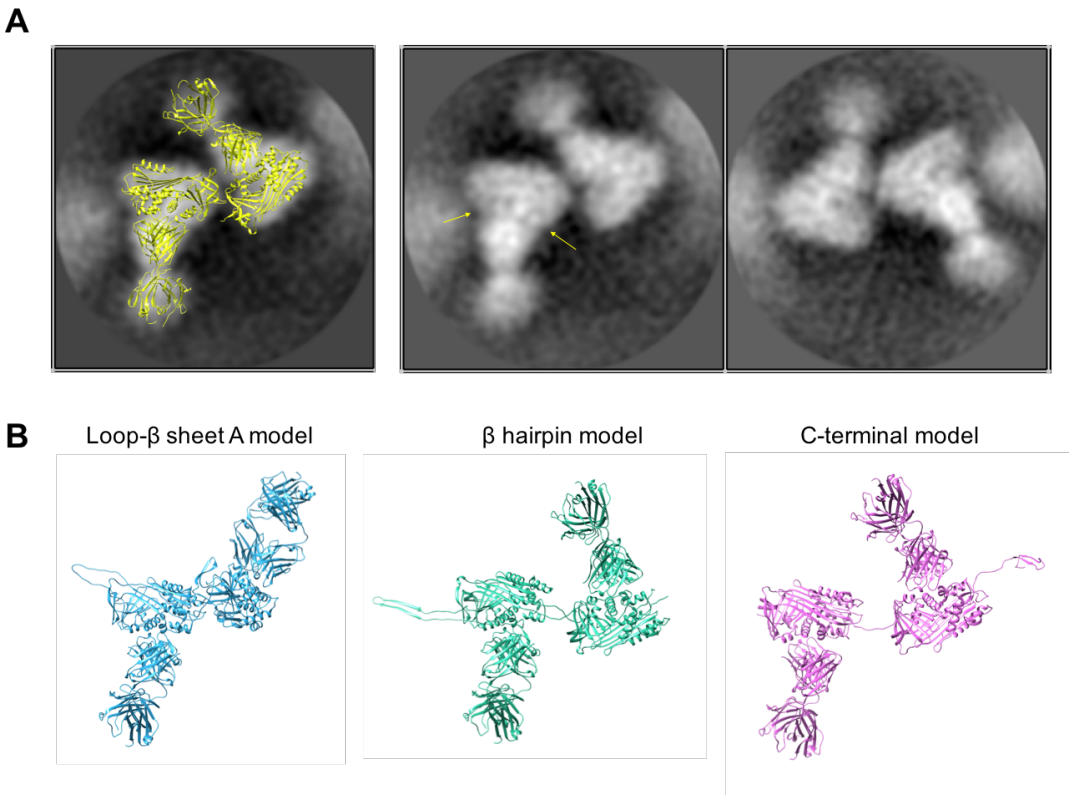


Figure 6.15. Cryo-EM class averages of dimers with inverted subunits. A) *AAT/Fab subunits (yellow) are superposed on one class average showing high shape complementarity (left). The class averages (right) show some internal details, highlighted by the arrows pointing to 3 helices in AAT. B) The three main polymerisation models were perturbed to have a -180° angle between subunits.*

Particles (93,000) were selected based on class averages that did not show such an extreme angle of rotation and used in a 3D classification job. The reference was made by averaging the two subunits from the previous model and fitting the averaged subunit back onto the model. Some of the 3D classes continued to display one nicely-formed subunit (Round 3, circled) but lacked the other. Images comprising these classes were selected and the 3D classification repeated several times (for example, Round 4, Figure 6.14). Model 3 in Round 4 was the closest to obtaining two resolved subunits but further 3D classification did not yield an improvement in the top subunit. The corresponding 15,000 particles were refined separately using either a dimer or a single monomer as references. The resulting reconstructions had a 0.99 correlation between them indicating that the second subunit was a product of the

particles and not the reference chosen. The reconstruction was soft-masked and the resolution calculated to be 17Å or 20Å for FSC values of 0.143 and 0.5, respectively (Figure 6.16A). Fitting copies of the AAT/Fab atomic model into both subunits with Chimera yielded measurements of 68Å distance and 31° between subunits.

Models 2 and 4 in Round 4 both displayed good fits of the atomic model to the bottom subunit; the corresponding particles were selected for 3D classification, where the good definition in the bottom subunit persisted and were refined against a monomer reference. An improved single subunit or monomer was obtained by masking out the other subunit. It had a clearer demarcation between the AAT and Fab within the subunit (Figure 6.16B). The structure had a resolution of 12Å or 15Å for FSC values of 0.143 and 0.5, respectively. However, internal details were still lacking.

Numerous other attempts were made to generate a sensible dimer model but only one subunit was ever defined. A single subunit was used as a reference in 3D classification rounds to avoid biasing towards a particular inter-subunit angle. However, no output models possessed a substantially formed top subunit. The 60° and 88° *ex vivo* NS-EM models were used as references in the 3D classification of the 108,000 particles to help identify corresponding particles and focus 3D reconstruction. Further selection of particles by 2D classification (to 40,000 particles) prior to selection 3D classification also did not alter the results. It was reasoned the lack of definition in the second subunit was due to a spectrum of angles between subunits in cryo-EM conditions.

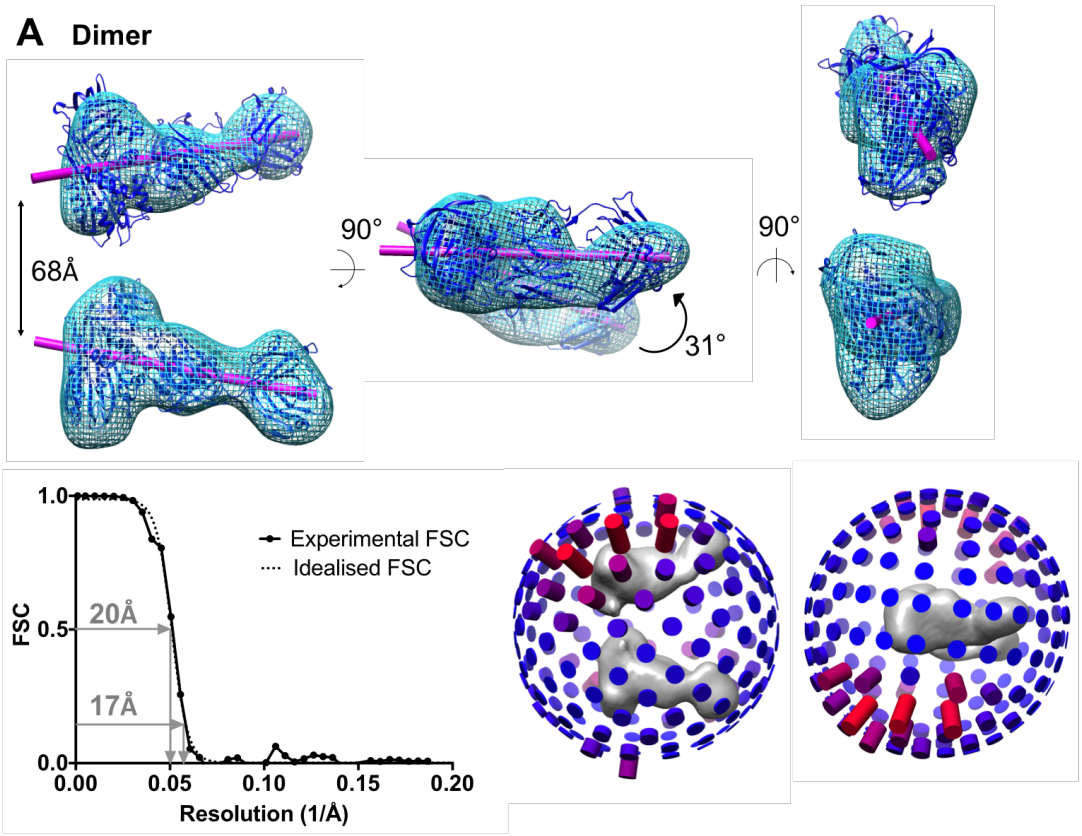

A Dimer
68Å
90°
31°
90°
1.0
0.5
0.0
FSC
20Å
17Å
Experimental FSC
Idealised FSC
0.00
0.05
0.10
0.15
0.20
Resolution (1/Å)
B Monomer from dimer
90°
90°
1.0
0.5
0.0
FSC
15Å
12Å
Experimental FSC
Idealised FSC
0.00
0.05
0.10
0.15
0.20
Resolution (1/Å)

Figure 6.16. Refinement of the dimer and single subunit from dimer particles. *The side, top and back views are shown are for the dimer reconstruction (light blue, top panel) of the 15,000 selected dimer particles. Atomic models of AAT/4B12 Fab (dark blue) are docked into the reconstruction with the central axes in pink used to make measurements. B) The same views are shown for the monomer reconstruction from 23,000 dimer particles, masking out the second subunit. The atomic model of AAT (pink) and 4B12 Fab (green) were fitted. FSC curves and plots of angular assignments are also shown for the reconstructions.*

6.3.10 Processing workflow for the AAT/Fab single subunit

The difficulty in obtaining a complete dimer reconstruction led to an alternative approach in determining the polymerisation mechanism: attempt to obtain a high-resolution structure of a single AAT/Fab subunit (~100kDa) from within the polymer. Reconstructing the single subunit would remove the problem of flexibility-derived heterogeneity between subunits, increasing the chances of achieving a higher resolution reconstruction. Resolving the path of the polypeptide chain would be sufficient to determine the strand swap between monomers within the polymer and which, if any, of the three polymerisation models are relevant to *ex vivo* material[32,78,86].

For the processing of the single subunit, two methods were proposed in parallel. The first was to extract monomers from already-picked dimers and the second was to pick monomers directly from the dataset. The first was not possible directly within Relion and was undertaken using Python scripts (written by Dr J. Irving) which transferred a translation centering the best-resolved subunit in 3D space to the 2D micrograph coordinates of the constituent particles. The purpose was to exploit the orientational information of the dimer. With simultaneous alignment and projection matching in the 3D classification step, effectively removing the orientational information from the dimer classification, the reconstructions all appeared quite messy and poorly-defined. When only projection matching was undertaken, which retained this prior angular information, the reconstruction had good definition

between the AAT and Fab. However, this approach again failed to provide the improvement in resolution sought after.

The second method was to pick single subunit particles from the beginning. A one subunit model was made by erasing density for the second subunit from a 3D dimer model created in the previous workflow. It was centred, filtered to 30Å and reprojected in uniform projections around the Euler sphere. The reprojections were used as references for autopicking and with a low-picking threshold, around 250,000 particles resulted. The particles were extracted into a 144-pixel (158Å) box and subjected to automated 2D classification and sorted several times to yield a final 63,000 of the best particles. Good class averages could be seen, with a range of side views but also including some top views. There were streaks apparent in the 2D class averages and these have been documented to result from overfitting[141]. The suggested approach is to limit the resolution in the E-step of the 2D classification algorithm[141]. Accordingly, the alignments were limited to 8Å and the background of particles varied from zeros to random noise (shown in Figure 6.17). The former appeared to have no effect, whereas the streakiness was reduced with the random noise background.

Particles selected solely on 2D classifications were subjected to a 3D refinement. The final structure fit the atomic AAT/Fab model very well in terms of shape complementarity and the structure refined to 9.8Å or 11.6Å with an FSC cut-off of 0.143 or 0.5, respectively. Particles were further sorted and selected by 3D classification. This led to a structure with better definition at the boundary between AAT and the Fab (marked by the arrow in Figure 6.17), compared to the refined 3D structure produced straight after 2D classification. Despite the greater definition, the resolution failed to improve (9.6Å or 12.8Å with an FSC cut-off of 0.143 or 0.5, respectively) most likely due to significantly fewer particles in the latter reconstruction. An angular distribution plot showed there were preferred orientations, which could also limit the resolution.

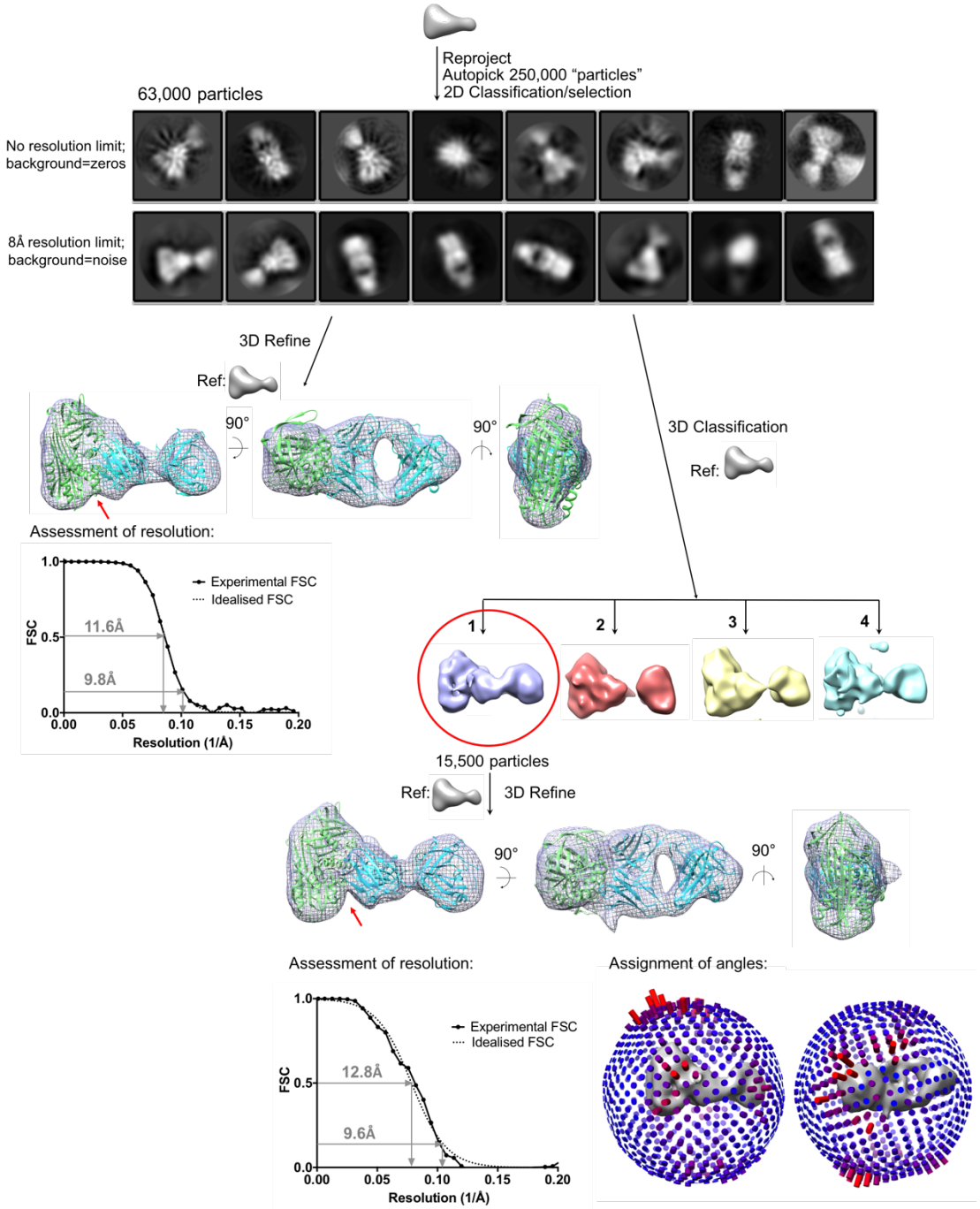


Figure 6.17. Cryo-EM workflow for image processing of the single subunit. *A single subunit was filtered to 30Å and reprojected to make references for autopicking. Particles were sorted by 2D selection to a final ~63,000. Different parameters (resolution limit and background setting) were trialled to remove streaks in the class averages. Particles were*

refined after 2D classification and also subjected to 3D classification for further particle section. After the latter route, refinement was carried out forming a model with greater definition around the binding point between AAT and 4B12 Fab. Reconstructions were thresholded to a volume corresponding to their molecular weight using Chimera (100kDa x 1.2Å^3/Da).

Internal details still could not be distinguished and so it was reasoned that the failure to reach higher resolutions was due to the misalignment of particles. Various parameters within the Relion 2D and 3D classification steps were changed to optimise alignment. This included the sampling search range and step of alignments, the regularisation parameter T value (determines relative weight of the experimental data), using a reference filtered to 20Å rather than 30Å, padding of images to a 300-pixel box (Figure 6.18A) and experimenting with masking in 2D and 3D (Figure 6.18B, C). Padding of images has been proposed to improve the accuracy of alignment as the Fourier transform is sampled more finely however it did not significantly alter the resolution in this case. Elliptical masking and hand-drawn masking of particles was attempted to remove the effect of neighbouring subunits on alignment, however the 2D masks caused the 3D reconstructions to exhibit spikes as a result of overfitting. Tighter 3D masks also did not improve alignments. Lastly, the particles were extracted from full-dose micrographs. Although they would have more electron-induced damage preventing the attainment of the highest resolution, the SNR would be higher and would aid alignment. Unfortunately, no improvement in the 3D reconstruction was observed either.

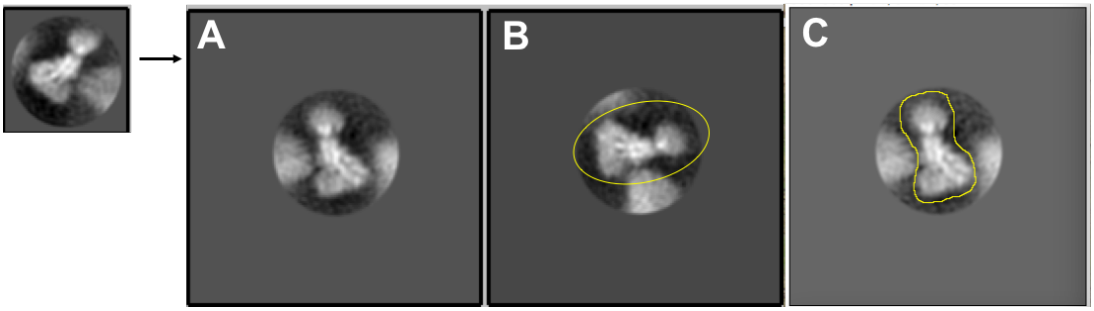


Figure 6.18. Approaches trialed to improve alignment of single subunit images. *On the left is an example class average in a 144-pixel box that had previously been used in reconstructions. A) shows a similar class padded into a 300-pixel box. B) shows a class with an outline of the elliptical mask applied and C) shows a hand drawn tighter mask drawn around the subunit to remove the influence of neighbouring subunits.*

6.3.11 Using the cryo-EM single subunit density to improve the AAT/4B12 Fab model and the polymerisation models

The refined cryo-EM single subunit structure was used to model the AAT and 4B12 Fab interaction using the same approach presented in 5.3.7. It was hoped the higher resolution of the reconstruction would make the modelling more reliable. The distribution of the scores of the fits of the 1000 models that were randomly perturbed and optimised into the density was again bi-modal (Figure 6.19A). Models with a low cross-correlation score had a visibly poor fit to the density. The top 15 models based on the approximate cross-correlation score reported by PyRosetta (score > 0.4) were analysed in Chimera. Three of these models were excluded as the AAT and Fab structures had collapsed into the centre of the EM density. All but one of the remaining 12 models showed a similar discrimination to that observed before, in which the Fab heavy and light chains favoured a particular orientation (Figure 6.19B). The exception was ranked as the 10th highest score. All top 12 models were in agreement with AAT in one orientation, indicating that the higher definition of the cryo-EM density provided greater confidence in comparison over the NS-EM density. The top scoring model was model 541 (Figure 6.19C). It showed a more feature-rich contour, including protruding density in the vicinity of a glycan that would be predicted to have a restricted degree of disorder due to steric hindrance.

Model 541 was also compared to the two top models from the modelling of the subunit interaction carried out in the NS-EM density (Figure 6.19D). AAT in models 329 and 787 was bound in two directions at almost 180° to one another, as this could not be distinguished based on the NS-EM density. Structural alignment was carried out in Chimera. Models 541 and 329 were more similar than models 541 and 787. In the former comparison, there is a small tilt of the Fab relative to AAT, such that in the cryo 541 model, the Fab is more perpendicular to AAT whereas in the latter, the Fab is at a drastically different angle relative to AAT. The equivalence of the 541 and 329 models is supported by the NS-EM density-fitting studies of the different polymer forms, where the 329-based models consistently had a higher score than the 787-based ones (Table 5.2).

The subunit structure represented by model 541 was used to create refined Fab-bound polymerisation models and repeat the fitting into the four new NS-EM densities made in Chapter 5. The C-terminal model of polymerisation had the highest cross-correlation score between the model and the density for all four structures (Table 6.2). Application of Dunn's non-parametric multiple comparisons test showed this was significantly different ($p < 0.0001$) to the β-hairpin and loop-β sheet A models for the equivalent liver and heat polymer comparisons carried out in 5.3.8. An example of the fits of the polymerisation models into the liver 60° structure is shown in Figure 6.20. The shape complementarity between the C-terminal model and the 60° structure is visibly high, with the Fab density providing important distance and geometry restraints. The linker also sits within the density. The next highest scoring model in this case is the β-hairpin model which also visibly fits well but the linkers are extended and out of the density. The worst scoring model is the loop-β sheet A model. Even with a minimally-inserted RCL, the linker is extended and appears to pull the second subunit out of the main area of density. Similar observations were made on analysis of the other reconstructions. These results most support the C-terminal model of polymerisation, for both liver and heat polymers. However, a mechanism of polymerisation cannot be definitively concluded in the absence of higher resolution information, which underpins the requirement for a new cryo-EM dataset.

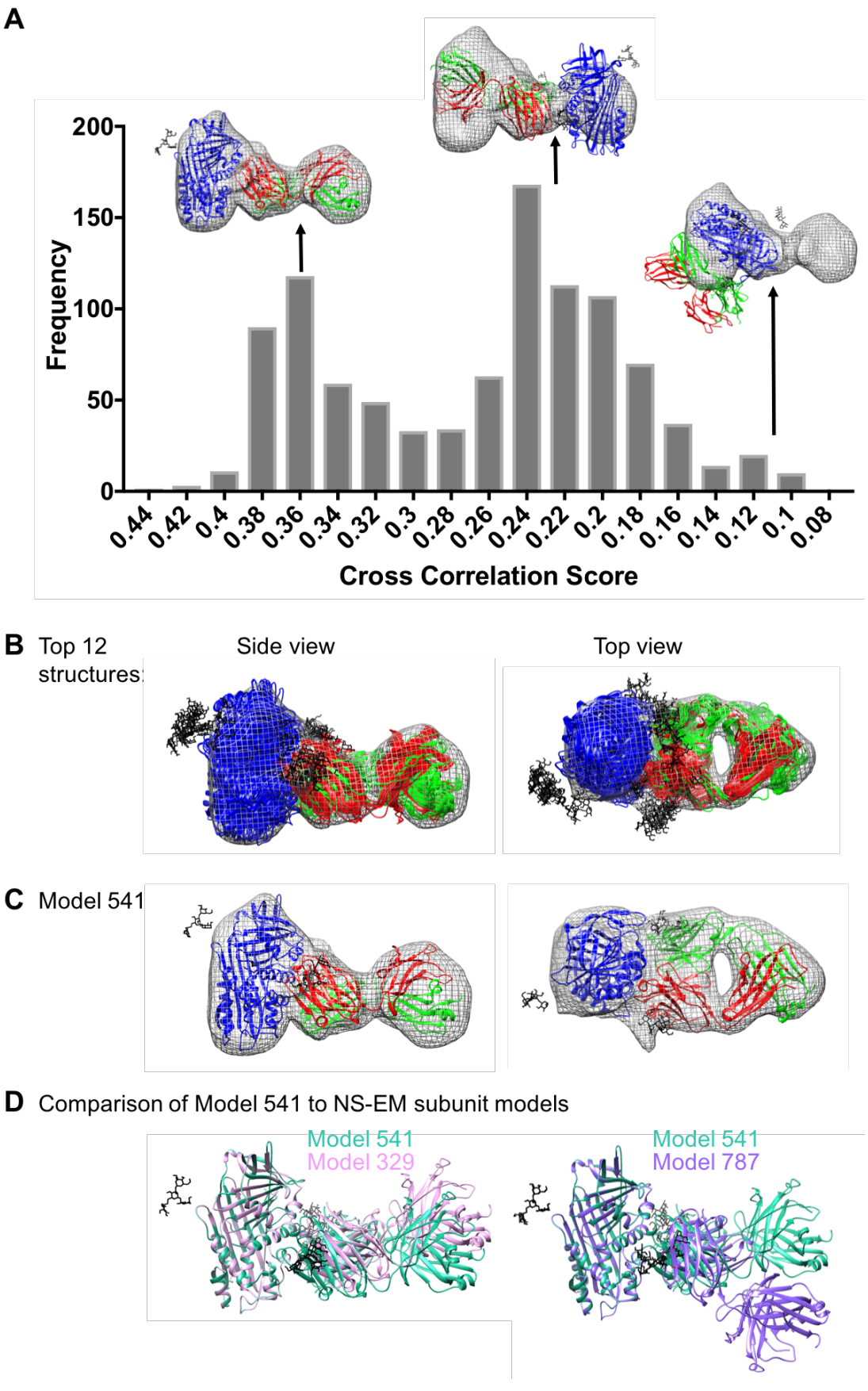

A
200
150
100
50
0
Frequency
0.44
0.42
0.4
0.38
0.36
0.34
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
Cross Correlation Score
B
Top 12 structures:
Side view
Top view
C
Model 541
D
Comparison of Model 541 to NS-EM subunit models
Model 541
Model 329
Model 541
Model 787

Figure 6.19. Improving the modelling of the 4B12 Fab/AAT interaction using cryo-EM-derived contour information. *A) 1000 random models were created and fit as a rigid body into the cryo-EM density of the single subunit. The cross-correlation score measured the fit to the electron density and the histogram shows the spread of model fits. Example models with differing scores are shown. B) The top 12 models with AAT in blue, the heavy chain in red and the light chain in green. Glycans are in black and the average density is thresholded at a volume corresponding to 1.5x the expected molecular weight. C) The highest-scoring model, Model 541. D) Alignment between Model 541 and the two highest-scoring models from the modelling into the NS-EM density (Model 329 and Model 787). Carried out in Chimera.*

		Loop-β sheet A model	β-hairpin model	C-terminal model
Liver polymer	**60°**	34.60	39.13	42.29
	88°	39.08	40.37	43.43
Heat polymer	**16°**	33.47	33.98	35.09
	59°	39.81	40.28	43.27

Table 6.2 Correlations of the fits of the updated polymerisation models using Model 541 into the NS-EM densities. *Modelling was carried out using PyRosetta and correlations were calculated between the models and the NS-EM structures. Values are the average of 10 models, all runs had a standard deviation of ≤ 0.02 and the highest score for each structure is highlighted in grey.*

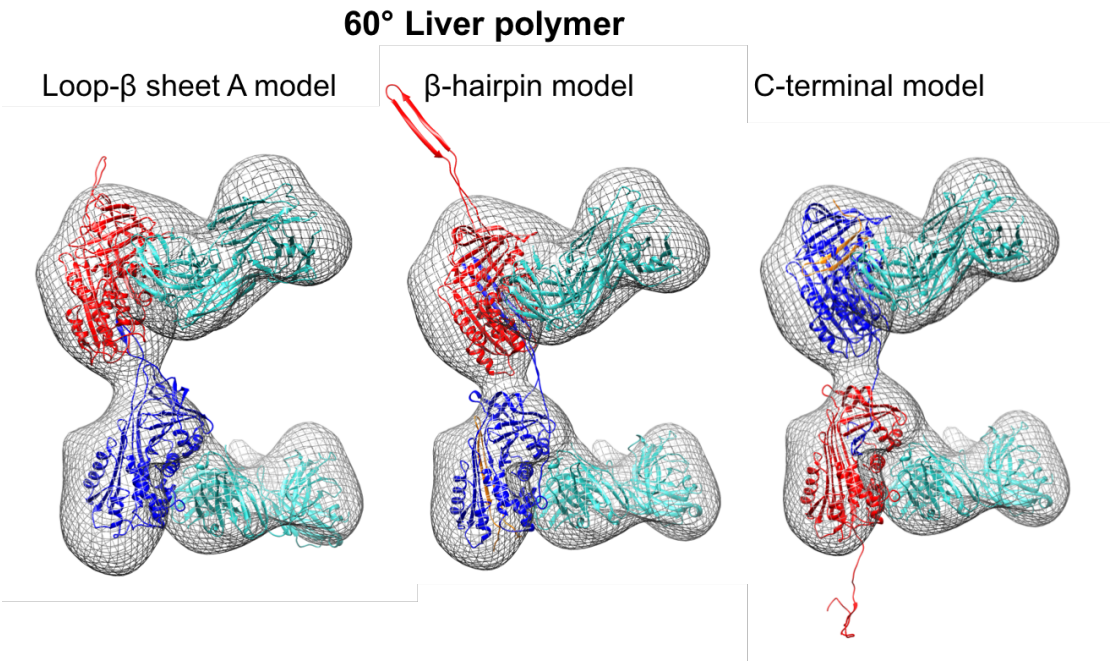


Figure 6.20. Fits of the cryo-influenced polymerisation models into the 60° liver polymer structure. *AAT subunits are shown in red, blue with the neighbouring AAT linkage in orange and the Fab subunits are in teal. Modelling was carried out using PyRosetta and visualised using Chimera.*

6.4 Discussion

6.4.1 Biochemical characterisation of liver polymers

Most research on the subject of AAT polymerisation has studied polymers that have been formed artificially[47,66,68,69,72]. Instead, this work aimed to study the structure of AAT polymers formed within a patient. This would provide the most physiologically relevant information towards understanding the polymerisation mechanism. There are three models of polymerisation (loop-β sheet A, β-hairpin and C-terminal models[32,78,86]) in contention, which differ in the linkage between neighbouring subunits. These are achievable *in vitro* but it is not known if they are relevant to the pathology of AATD.

AAT polymers were isolated from liver samples of a 46-year-old ZZ male who had undergone an orthotopic liver transplant as treatment for the liver disease associated with AATD. The method of polymer extraction from the liver tissue avoided the use of reagents that are known to affect AAT polymerisation and instead used sonication to dissociate IBs and release the soluble polymer[32,280,286]. Due to the extreme stability of polymers[47,67], it is believed that this method of extraction does not alter polymer structure (i.e. the linkage), however there is the possibility that sonication dissociates the polymers into smaller chains[287].

Separation of the components of the tissue extracted by denaturing gel showed the presence of a dominant protein band at the size of AAT in the liver sample and the identity was confirmed by western blot. The sample was contaminated with other proteins present at much lower concentrations (Figure 6.2). The identity of most of these was not established, except the second dominant band of carbamoyl phosphate synthetase 1 (CPS1) which was identified by proteomic analysis. CPS1 is involved in the production of urea. An association between this protein and AATD has not been studied, nor has there been a link established between the urea cycle and AAT.

The liver polymers were found to have a high mannose glycosylation pattern in contrast to plasma M-AAT, which has mature glycans. Carbohydrate analysis of AAT

isolated from IBs has previously indicated AAT to be of the high mannose form which is susceptible to EndoH cleavage[281,286,288]. The glycosylation pattern gave confidence that the polymers originated in the ER and were suitable for structural characterisation. Immature glycosylation also explained the faster migration of liver Z-AAT on a denaturing gel compared to M-AAT as complete deglycosylation of both ZZ liver and M plasma monomer using PNGaseF removed the migration difference (Figure 6.5).

A non-denaturing western blot gel (and NS-EM) showed that all the AAT extracted from the liver sample was in polymer form and that the polymers were recognised by 2C1. This shows the sonication was not sufficiently harsh to dissociate chains into monomer. The mass of polymers isolated was quantified by ELISA. Dissociation of the IB preparation yielded a total of 3.2 mg of polymer, representing the yield from 3.9 g of liver. The insoluble material, pelleted after sonication, was re-sonicated a total of four times until complete depletion of AAT was achieved. Only small amounts of AAT could be seen after each repeated sonication (not shown) indicating that the majority of the polymer had been released after one sonication round. While not fully quantitative, the quantities recovered can give some indication of polymer burden. It is likely therefore 0.8 mg polymer/g liver is a slight underestimate for this ZZ patient. This is relatively similar to polymer extraction from another ZZ homozygote where 0.9 mg polymer/g liver was recovered. Furthermore, isolation from two MZ heterozygotes yielded 0.6 and 0.2 mg polymer/g liver. Polymer levels from liver samples have not been quantified before but large heterogeneity between clinical manifestations of patients has already been documented[102,289].

6.4.2 Cryo-EM studies required further purification of liver polymers

The isolated liver polymers were the major component of the sample but with protein and possibly other impurities that could not be detected by the methods used. Although EM is one of the few structural techniques that can mitigate the effect of heterogeneity and impurities, high sample purity aids with grid preparation as well as particle picking[290]. This was the first time that liver polymers had been studied using cryo-EM. The potential effect of the impurities on freezing was proposed whilst trying

to optimise cryo-EM grids. Ice thickness and sample concentration across the grids was highly variable. Furthermore, non-ice crystals were observed. The composition of these were unknown but the crystals could contain other protein contaminants or components from membrane debris including lipids.

Previous attempts to reduce the level of impurities using a commercial antibody-based $Alpha_1$-Antitrypsin Select column and a custom-made rabbit polyclonal antibody-conjugated resin were unsuccessful. This may have been due to properties of the column (conjugated antibody fragments not available to bind) or detection methods insensitive to the low amounts of material available for trial. Access to more patient material and a new chromatography system enabled another purification method to be trialled: the use of anion exchange chromatography. As with monomeric AAT, the liver AAT polymers bound tightly to the Q-sepharose column and were eluted in the salt gradient. The UV trace of the elution indicated the removal of many contaminants, including material able to absorb light at 280nm but unable to be detected by the Coomassie protein stain. The increased purity was clear from NS-EM (Figures 6.11 and 6.12).

Following this successful chromatography step, a Q-sepharose column was also used to purify away excess 4B12 Fab following the Fab-labelling step. At pH 8.0, 4B12 Fab eluted at a lower salt concentration, suggesting a less overall negative charge, and so could easily be separated from the polymer. The salt concentration at which labelled polymers were eluted barely changed relative to unlabelled liver polymer, indicating that 4B12 did not obscure any major negatively charged surface areas of AAT; 4B12 binds to helix A and I in an area which is predominantly neutral in charge and so if anything increases the strength of binding to the column[117].

After further purification, freezing in the subsequent optimisation process resulted in cryo-EM grids of more consistent ice thickness, highlighting the effect that impurities can have on freezing.

6.4.3 Additional differences identified between liver polymers and heat polymers

Structural comparisons were made between heat-induced and liver polymers visualised by NS-EM in Chapter 5, however the use of cryo-EM highlighted a few other characteristics. As suggested previously, longer linear polymers were apparent in the NS-EM micrographs of the *ex vivo* polymer sample compared to the heat polymers. These *ex vivo* polymers are inconsistent with a SAXS study that characterised the solution structures of short-chain Z-AAT polymers formed by Pichia Pastoris[95]: the authors found that the trimer, tetramer and pentamer polymers were circularised, and proposed to therefore support the circular C-terminal model[86]. The observation of longer *ex vivo* polymers are supported by a non-denaturing western blot showing higher molecular weight protein in the liver sample (Figure 6.3). The isolation of longer polymers may support the notion that higher-order polymers are sequestered into IBs as a cellular protective mechanism whereas smaller polymers are secreted and/or degraded[105,193]. The western blot also highlights the differing appearance of the polymers on a denaturing gel; the heat polymers show a clear laddering pattern whereas the liver polymers appear as more of an indistinct 'smear' with low-molecular weight laddering at different positions to the heat polymer. As non-denaturing gels separate on the basis of charge as well as size, the immature glycosylation of AAT that has not left the ER may result in an altered running pattern on the gel and heterogeneity of glycosylation or other post-translational changes may contribute to the smear[291]. Notably, polymers of non-glycosylated recombinant AAT also resolve as a smear, possibly due to higher-order interactions between individual polymer chains[117].

Whilst trying to find a suitable purification scheme for the *ex vivo* polymers, heat-induced polymers of plasma AAT were used as a surrogate to conserve material. Both polymer types were eluted from a Q-sepharose column, using the same buffer. The peaks of their elution differed somewhat and were at 300mM and 340mM salt concentration for the liver polymer and heat polymer, respectively. This suggests that the major polymer species in each have slightly different overall charges, with the liver polymers being less negatively charged. This is a comparatively small difference

compared to that between monomer (240mM NaCl) and polymer, and so is unlikely to be because of a markedly different conformation and is insufficient to support a difference in inter-subunit linkage. It is also probably not due to polymer length as polymers largely elute under narrow conditions in comparison to a broad size profile. It is most likely due to differences in glycosylation again, as the sialic acid sugar residues on the mature glycans of the heat polymer are negatively charged causing them to bind tighter to the column. This corroborates the finding that the isoelectric point of plasma AAT is more anodal than liver AAT[281].

Lastly, optimum cryo-EM freezing conditions differed between the two samples (10 blot force, 3 sec blot for the liver polymer and 5 blot force, 4 sec blot for the heat polymer). This may be due to different behaviours of the sample in solution and with the grid (reflecting differences in sample including different glycosylation states) or due to the presence of other components in the liver polymer sample. Optimum freezing conditions are found by trial and error and vary between proteins, so it is difficult to draw any conclusions about differences between the two polymer species[290].

6.4.4 The loop-β sheet A model is not consistent with the extreme angle classes

Dimer particles from the cryo-EM dataset were processed using Relion[122]. Relion was chosen after proving successful in producing well-defined reconstructions and dealing with heterogeneity in the NS-EM work. Particles were automatically picked using projections from a dimer model reconstructed from a limited subset of the cryo-EM data, based on the cryo-EM data, which had a 27° angle between subunits. Sorting of the particles using 2D classification showed most particles with well-resolved Fab domains appeared to have an acute angle between subunits, however two classes had an extreme angle – nearer 180° – between subunits (Figure 6.15A). Such an angle had been observed previously in the NS-EM heat polymer work (Figure 5.8C). The ability to detect these particles within the data gave confidence that the autopicking references and parameters were sufficient to detect relevant particles that differed significantly from the references.

The classes with the extreme angle between subunits appeared to be of higher resolution than other classes, with possible elements of secondary structure identifiable. However, they represented only a very small proportion (2.5%) of the dataset and thus were insufficient to proceed to 3D reconstruction. Furthermore, identifying equivalent particles with the Fab domains orthogonal to the plane of the grid was likely to prove extremely challenging. The inability to manually 'seed' multiple distinct classes in Relion provides an impediment to this kind of analysis. Nonetheless, the classes definitively showed that the polymerisation model would have to be consistent with a complete flip of the AAT chain. Modelling studies showed that to maintain the inter-subunit linkage in the loop-β sheet A model of polymerisation, for a dimer with a rotation of 180° between subunits, one subunit had to be rotated out-of-plane (Figure 6.15B). This would create an end view of one AAT/Fab subunit. However, the extreme angle classes clearly represent a dimer with side views of both subunits. Therefore, the version of the loop-β sheet A model – with an insertion of P3-P8 of the RCL – is not compatible with this conformation of dimer and is unlikely to represent the linkage of polymers formed *in vivo*. Both the β-hairpin and C-terminal models could yield a dimer conformation with side views of both subunits in a 2D projection.

6.4.5 There is a lack of a well-defined cryo-EM dimer reconstruction

Cryo-EM is the only structural technique that has the capability to obtain high-medium resolutions of a flexible, heterogeneous sample such as the AAT polymers[259,292,293]. However, conformational flexibility is still a major hurdle to overcome when attempting to obtain a complete reconstruction at a resolution sufficient to discern internal protein features. This is most likely the case for the labelled dimer reconstruction.

When working with the initial set of 16,000 particles, which were selected based on their good definition after the first round of autopicking, a labelled dimer reconstruction was obtained but one subunit of the dimer was always more poorly formed than the other. The definition was improved by updating the reference model for the 3D classification job (created by duplicating the well-formed subunit of an

output model on the other subunit), but the definition was still unequal between subunits. At this stage, it was postulated that this could be due to the flexible linker between subunits or low particle number in the reconstruction. After re-picking and sorting the dataset to yield the final 93,000 small-angle dimers, 3D classification models continued to lack a top subunit with a resolved Fab domain. Further selection rounds and classification against a monomer and other dimer models offered no improvements to this region. The best dimer reconstruction is shown in Figure 6.16A with a 31° angle between subunits and refined to 20Å. As the structure was obtained using a monomer reference, this dimer conformation is likely to be representative of a subset of the dataset rather than being the result of a biased input reference. However, the heterogeneity observed in the 2D classes and the poorly defined second subunit, especially in other models, may highlight a continuum of angles that is present between subunits in the native conditions of cryo-EM. A reconstruction from particles corresponding to dimers of slightly different rotations would be expected to lead to averaging out and degradation of subunit definition. In comparison, the application of NS may have rigidified the polymer and favoured the angles observed in the reconstructions presented in Chapter 5. To obtain a better reconstruction of the dimer, more particles with narrower classification will be needed to reduce the effect of the flexibility. Most reconstructions solved to resolutions of 3-4Å are obtained from between 30,000-500,000+ particles[294]. The recent "multi-body refinement" function incorporated into Relion v3.0, which separates structures into rigid bodies for refinement, may also help generate improved reconstructions and provide insight into the principal conformational movements in the dimer[295].

The challenges imposed by polymer flexibility was anticipated for the image processing. Multiple approaches were taken to polymer preparation in order to optimise its characteristics for imaging and data processing. Alternate labelling strategies focused on decreasing the linearity of the polymers and/or introducing extra asymmetry. Both alterations would be expected to aid alignment. Fabs from different mAbs were bound to heat-induced polymers, but the flexibility of chains was unchanged. This was not surprising for mAbs that recognise AAT monomer, as one corresponding Fab molecule would likely bind one AAT molecule and not influence the flexibility between neighbouring subunits. However, there was a chance that

mAbs that only bound polymers, such as 1A3, could have a binding site that bridged two AAT subunits and consequently would reduce dimer flexibility. Polymers were also bound with 4B12 Fab and 9C5 Fab simultaneously. Both these antibodies can bind to a monomeric AAT at the same time and can bind to polymeric AAT so each subunit of the polymer was bound by both Fabs[116]. The linearity of the polymers was unchanged, but their complex decoration was apparent (Figure 6.8). It was difficult to infer the relative binding sites of the two Fabs on the AAT subunit, but they did not appear to bind diametrically opposite each other. It was also hard to know whether having two Fabs bound would help or hinder alignment as the two Fabs would appear similar in projections. Therefore, as the addition of 9C5 Fab did not decrease flexibility, this strategy was not pursued further. The binding of 4B12 Fab_2 fragments or 4B12 mAb to polymers also did not influence flexibility. Both have two antigen-binding sites and could bind and potentially generate a geometrically constrained 4B12-AAT dimer subunit within a polymer. However, these methods appeared to increase heterogeneity, with Fab_2 and mAb bridging between two polymers.

A more complicated approach to extend the polymers on a grid that could be trialled in future work has been applied to the flexible titin molecule[296]. Titin molecules are long single polypeptides made up of multiple domains which have been observed to adopt varying conformations from a straight structure to a random coil. Extension of the molecules or 'molecular combing' as visualised by EM was achieved using centrifugation and liquid de-wetting steps. This produced shear flow and surface tension forces respectively, that resulted in a straightening of the molecules. Novel methods could also be developed that utilise linear scaffolds such as cationic polyelectrolytes to straighten the polymers.

6.4.6 The cryo-EM single subunit allows refinement of the 4B12 Fab/AAT interaction

The computational method taken to remove inter-subunit flexibility-derived heterogeneity was to focus on the structure of the 4B12 Fab/AAT single subunit from within polymer chain. At a high enough resolution (~8Å), the polymer linkage, which differs between the models of polymerisation, could be determined. Areas of

difference between the models would be at the top of β-sheet A to see whether the RCL was self-inserting (C-terminal model) and where the linkers leave and enter the main AAT body.

A single subunit structure was determined to 12.8Å resolution. This provided the first observation of the characteristic 'hole' between the constant and variable domains of the Fab in the reconstructions. The better definition of the reconstruction allowed the 4B12 Fab and AAT interaction to be refined. The cryo structure – unlike the NS-EM structure – could unambiguously determine the direction of AAT subunit and potentially revealed density for some of the AAT glycans. Cryo-EM is the most suitable structural method to resolve untrimmed glycans, and many have been observed before, particularly on viruses[297]. A higher resolution structure that resolved the glycans would be the first of AAT in glycosylated form.

The tilt between the interacting Fab and AAT slightly changed between the cryo-EM 541 model and the NS-EM 329 model. It has been documented that Fabs can bind to their antigens with a degree of flexibility and so this is likely to represent a variation in binding in different conditions[226,298]. Unfortunately, at the resolution obtained, no internal details that pointed towards a polymerisation model could be observed. It is likely the resolution was limited by poor alignment of particles; Relion estimated that particles could only be assigned accurate translations and rotations up to 3 pixels and 8 degrees, respectively. Difficulties with alignment is further supported by the radially extending streaks in the class averages. The low SNR of particles small in size would be a significant contributor to this, which again highlights the use of the 4B12 Fab as a tool for doubling the AAT molecular weight for EM studies[262]. Furthermore, implicit in the reconstruction algorithms is the notion that the particle of interest is surrounded by solvent. However, as the subunit occurs sandwiched between density provided by adjacent subunits, this distinction between solvent and protein would be blurred somewhat. Attempts to improve the alignment within Relion did not work suggesting that the either the number of particles were limiting for a complex this size and/or the images were not of high enough quality to aid higher alignment (e.g. particles in thicker ice have higher background).

6.4.7 Approaches for improvement in the next cryo-EM dataset

The preliminary work carried out has highlighted the challenges with structural reconstructions of the AAT polymer but also that it shows promise of feasibility. For example, reconstruction of the single subunit produced significant improvements in definition relative to the dimer. Application of new developments such as multi-body refinement and PCA to reveal motions in the sample may also prove useful[295]. Another dataset with more and better quality images will be essential to obtaining high resolution reconstructions. This is especially important when solving structures of small (100 kDa) complexes. Datasets that have obtained high resolution (greater than 4Å) for small proteins (< 150 kDa) have consisted of 1500-2500 micrographs, with greater than 500,000 particles narrowed down to the top 10% (50,000) of particles that constitute the final reconstruction[127]. The 3.2Å reconstruction of haemoglobin, which is 64 kDa, was generated from phase plate data and consisted of more than 175,000 particles[299]. The dataset of the *ex vivo* AAT polymer had just over 200 micrographs, with 108,000/63,000 particles, narrowed down to 15,000/15,500 particles in the final dimer/monomer reconstructions, respectively. Separation of particles in groups based on heterogeneity had to be compromised with the number of particles in a structure. With a bigger dataset, it is hoped that higher resolution structures of the single subunit, the near planar dimer and the minor but well-defined species of the opposing-angle dimer will be obtained. Areas for improvement include sample preparation, grid preparation and data collection set-up.

The sample has been efficiently purified from the liver tissue through two chromatography columns. However, polymers could be further purified according to size through a gel filtration column, but this may not directly influence attainment of a high-resolution structure. The most challenging feature of the sample is the flexibility; an approach to reduce this could have dramatic effects on reconstruction. As the polymers originate from a human, it is not possible to alter the fundamental basis of the polymerisation process to create a more favourable sample for structure studies. The existing antibody toolkit could prove useful but for different approaches to be screened, sample amounts need to be sufficient.

There is a wide range of variations in grid preparation that could result in improvements in data quality. Issues apparent in the preliminary dataset included areas of high polymer concentration and sample aggregation that had to be discounted from analysis. A higher dilution of sample should be applied to grids and the sample buffer could be adapted to contain a higher salt concentration (up to 500mM), which may reduce inter-polymer interactions. Concentrations of detergent such as Tween20, below its critical micelle concentration, may also facilitate spreading of the sample[300]. Linked to these points is the use of graphene oxide as a support layer on grids[301]. This will enable less sample to be used per grid, which is important given the small sample amounts available and would limit the dimensionality of projections. Most importantly, the ice thickness should be homogeneous in thickness to maximise the SNR from a small particle[290]. A lower blot force than the one used in this data collection on the Vitrobot is now recommended by most users, with optimisation primarily achieved by varying blotting time (personal communication, Sample Preparation course at Diamond, Oxford).

Lastly, changes in the data collection set-up could prove essential in obtaining high resolution. Only a relatively small dataset was collected (200 micrographs). Routine data collection can now collect 2000+ micrographs. This would significantly increase the number of particles used for reconstruction. Furthermore, the use of a phase plate would aid in visualisation and alignment of small particles[302].

6.4.8 Future work

After finalising the *ex vivo* polymer structure by cryo-EM, the logical next step would be follow the same process for the heat-induced polymer to reach a conclusion about their structural similarity at high-resolution. Significant competition for the available cryo-EM data collection time notwithstanding, use of such a sample could permit a wider exploration and optimisation of data collection parameters due to the ready availability of material. A longer-term goal is to determine 3D reconstructions of other *ex vivo* AAT polymer types to identify any structural differences in the linkage. These include other pathological mutations within AAT, for example the S mutation which is located on helix G, on the opposite face of β-sheet A and the Z mutation[303].

Furthermore, it has not been shown whether polymers from different serpinopathies have the same molecular structure. The generality of the polymerisation mechanism can be directly addressed by studying the structure of other serpin polymers, such as mutant neuroserpin, which is the basis for a rare form of hereditary dementia. Similar approaches to sample treatment and optimisation (e.g. binding neuroserpin-specific Fabs) could be applied to aid the EM structural determination. Polymerisation of neuroserpin has been studied biochemically and biophysically and shows similarity to the process observed with AAT: the process proceeds via an intermediate conformation and appears to occur as a result of a destabilised native state. In support of this, the H334D Kings mutation in AAT is homologous to a mutant of neuroserpin that polymerises in FENIB[76]. However, neuroserpin also readily forms the latent conformation, an inactive monomeric hyperstable state that AAT does not[24,26]. Thus, there are differences and it is unknown whether the polymerisation mechanism and the polymeric linkage are common to all serpins.

7.0 Conclusions and outlook

7.1 The importance of studying AAT polymerisation

Point mutations within the SERPINA1 gene, which encodes alpha$_1$-antitrypsin (AAT), can lead to polymerisation of the protein *in vivo*. Polymer formation and retention in hepatocytes, the main site of their synthesis, predisposes individuals with AAT deficiency to liver cirrhosis[42]. Local polymerisation in the lungs, as well as the deficiency of functional circulating AAT, predisposes patients to emphysema[42]. The structure of the *in vivo* polymer, and in particular the inter-subunit linkage, is unknown. As almost all biophysical and biochemical studies of serpins to date have made use of *in vitro* polymers, there is an unresolved uncertainty as to what extent these data can be extrapolated to *in vivo* polymers[32,78,86]. It has already been shown that one type of polymer, produced using denaturant, is unlikely to be representative of pathological polymers[69,76,117]. This lack of understanding prevents a fully described polymerisation mechanism and is of direct relevance to attempts to develop therapeutic agents to prevent polymerisation in patients and develop diagnostic tools for quantifying polymer burden in an individual.

This work has aimed to structurally characterise isolated pathological AAT polymers and compare such polymers to those formed by heating. This has been achieved through biochemical techniques, monoclonal antibodies and electron microscopy.

7.2 AAT-specific monoclonal antibodies are useful investigative tools

Over 30 novel AAT-specific mAbs were developed in a hybridoma screen. Many showed interesting conformational selectivity against AAT conformers but 1D9 stood out as the only, and first, monomer-selective antibody. Its preference for monomer over polymer, and definition of its epitope in atomic detail, provide further evidence of conformational changes that occur on polymerisation. From the 1D9 epitope, one such area of change is the RCL and s1C of AAT, which is predicted by all polymerisation models[32,78,86]. 1D9 also blocked heat-induced polymerisation and inhibitory activity of AAT against chymotrypsin, highlighting the importance of the

RCL/s1C to these mechanisms. There is general interest in the field in developing better tools to define biomarkers that predict disease severity, progression and outcome. The monomer-selectivity of 1D9 lends itself to such purposes, and raises the possibility that it may be suitable for applications in diagnosing AATD by accurately measuring monomer levels in patient plasma. The conditions to carry out this quantification, and the association between the levels of native AAT in circulation and disease, will now need to be investigated.

Other mAbs of interest that require further characterisation in future projects are the 2H2 and 1A3 antibodies. These show M-AAT specificity and polymer-specificity, respectively. In particular, the 2H2 mAb/Fab will be useful in studies investigating polymers from heterozygous MZ-AAT patients[304]. The previously-developed 4B12 mAb has also been instrumental to the EM studies by providing extra signal and asymmetry in image processing. Clever engineering of AAT-specific mAbs could further aid sample behaviour, in particularly flexibility for cryo-EM work.

7.3 Heat polymers and liver polymers have more similarities than differences

In general, this work has supported the notion that heat-induced polymers and *ex vivo* polymers are structurally similar (based on 2C1 immunoreactivity) and therefore heat-induced polymers are a valid model for AATD research[76]. Qualitatively, heat-induced polymers and *ex vivo* polymers appeared similar in NS-EM micrographs in terms of their flexibility, shape with bead-like subunits, and consistent labelling by 4B12 Fab. The NS-EM reconstructions of both polymer types also had similar inter-subunit distances and angular relationships and the fitting of the polymerisation models yielded similar results. Slight differences between the two polymer types included different proportions of linear and circular polymers, the length of polymers, differing charges (as evidenced by migration on non-denaturing PAGE and binding to anion-exchange resin) and a lack of equivalent heat polymer reconstruction with a subunit angle near to 90°. Most of these differences do not necessarily indicate an alternate polymer linkage. The most definitive evidence for heat-induced polymerisation and *in vivo* polymerisation resulting in different polymer structures is the recognition of *ex vivo* polymers, but not heat-induced polymers, by 1D9. This is

especially important considering the involvement of the RCL/s1C (the 1D9 epitope) in polymerisation. To definitively assess the validity of heat polymers as a model for *ex vivo* polymers, cryo-EM studies of the heat-induced polymer will need to be carried out.

7.4 The C-terminal model of polymerisation is most consistent with *ex vivo* data

Out of the three major polymerisation models, the C-terminal model was the highest-scoring and the best-fitting into the NS-EM reconstructions[86]. It was also able to accommodate the near-anti-parallel arrangement of the AAT chain observed in a few of the well-defined cryo-EM 2D dimer classes. Furthermore, a comparative structural analysis between the 1D9-AAT complex and the models of polymerisation showed that the main impediment to binding would be steric clashes with neighbouring AAT subunits. The C-terminal model of polymerisation, with a single peptide linker and therefore the greatest flexibility, would be the most likely to accommodate the distortion of the chain required for recognition. Therefore, collectively these data support the C-terminal model of polymerisation over the loop-β sheet A and β-hairpin models. The initial cryo-EM studies have shown the feasibility in obtaining AAT density and resulted in the refinement of the single subunit model between AAT and 4B12 Fab. Now that the purification protocol has been established, with a larger and higher quality cryo-EM dataset, with an increase in the number of images, optimisation of the concentration and spread of the polymers and ice thickness, it is hoped the nature and extent of intermolecular interactions in patient-derived polymers can be definitively determined from a dimer and/or single subunit structures. There is also the possibility of investigating alternative methods to limit polymer flexibility from the greater access to patient material. High-resolution electron density could reveal a new model of polymerisation, which could not be predicted at the resolution achieved in this work. Knowledge of the degree to which the AAT subunit structure is perturbed would be expected to yield insight into whether polymerisation proceeds from a native-like or substantially unfolded intermediate and predict testable ways in which polymerisation can be altered *in vitro* or in a cell model of disease. Lastly, once a cryo-EM workflow has been established, it will be of clinical interest to determine a structure of *ex vivo* polymers isolated from

a MZ-AAT patient, patients with other AATD-mutations in other regions of the AAT molecule and of other serpins, for example neuroserpin. This would address the question of whether the polymerisation mechanism is applicable to a variety of mutations and serpins. Ultimately, this could extend the relevance of structure-based drug design in therapeutic research from AATD to other serpinopathies.

8.0 Bibliography

1. Huntington JA. Serpin structure, function and dysfunction. *J Thromb Haemost.* 2011;9 Suppl 1:26-34.

2. Gooptu B, Lomas DA. Conformational pathology of the serpins: themes, variations, and therapeutic strategies. *Annu Rev Biochem.* 2009;78:147-176.

3. Gettins PG. Serpin structure, mechanism, and function. *Chem Rev.* 2002;102(12):4751-4804.

4. Law RH, Zhang Q, McGowan S, et al. An overview of the serpin superfamily. *Genome Biol.* 2006;7(5):216.

5. McGowan S, Buckle AM, Irving JA, et al. X-ray crystal structure of MENT: evidence for functional loop-sheet polymers in chromatin condensation. *EMBO J.* 2006;25(13):3144-3155.

6. Gardill BR, Vogl MR, Lin HY, Hammond GL, Muller YA. Corticosteroid-binding globulin: structure-function implications from species differences. *PLoS One.* 2012;7(12):e52759.

7. Knaupp AS, Bottomley SP. Serpin polymerization and its role in disease--the molecular basis of alpha1-antitrypsin deficiency. *IUBMB Life.* 2009;61(1):1-5.

8. Elliott PR, Pei XY, Dafforn TR, Lomas DA. Topography of a 2.0 A structure of alpha1-antitrypsin reveals targets for rational drug design to prevent conformational disease. *Protein Sci.* 2000;9(7):1274-1281.

9. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605-1612.

10. Anfinsen CB. Principles that govern the folding of protein chains. *Science.* 1973;181(4096):223-230.

11. Seo EJ, Im H, Maeng JS, Kim KE, Yu MH. Distribution of the native strain in human alpha 1-antitrypsin and its association with protease inhibitor function. *J Biol Chem.* 2000;275(22):16904-16909.

12. Boudier C, Bieth JG. The reaction of serpins with proteinases involves important enthalpy changes. *Biochemistry.* 2001;40(33):9962-9967.

13. Huntington JA, Read RJ, Carrell RW. Structure of a serpin-protease complex shows inhibition by deformation. *Nature.* 2000;407(6806):923-926.

14. Loebermann H, Tokuoka R, Deisenhofer J, Huber R. Human alpha 1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J Mol Biol.* 1984;177(3):531-557.

15. Lawrence DA, Ginsburg D, Day DE, et al. Serpin-protease complexes are trapped as stable acyl-enzyme intermediates. *J Biol Chem.* 1995;270(43):25309-25312.

16. Wright HT, Scarsdale JN. Structural basis for serpin inhibitor activity. *Proteins.* 1995;22(3):210-225.

17. Stratikos E, Gettins PG. Formation of the covalent serpin-proteinase complex involves translocation of the proteinase by more than 70 A and full insertion of the reactive center loop into beta-sheet A. *Proc Natl Acad Sci U S A.* 1999;96(9):4808-4813.

18. Herve M, Ghelis C. Conformational stability of the covalent complex between elastase and alpha 1-proteinase inhibitor. *Arch Biochem Biophys.* 1991;285(1):142-146.

19. Kaslik G, Patthy A, Balint M, Graf L. Trypsin complexed with alpha 1-proteinase inhibitor has an increased structural flexibility. *FEBS Lett.* 1995;370(3):179-183.

20. Dementiev A, Dobo J, Gettins PG. Active site distortion is sufficient for proteinase inhibition by serpins: structure of the covalent complex of alpha1-proteinase inhibitor with porcine pancreatic elastase. *J Biol Chem.* 2006;281(6):3452-3457.

21. Kounnas MZ, Church FC, Argraves WS, Strickland DK. Cellular internalization and degradation of antithrombin III-thrombin, heparin cofactor II-thrombin, and alpha 1-antitrypsin-trypsin complexes is mediated by the low density lipoprotein receptor-related protein. *J Biol Chem.* 1996;271(11):6523-6529.

22. Dementiev A, Simonovic M, Volz K, Gettins PG. Canonical inhibitor-like interactions explain reactivity of alpha1-proteinase inhibitor Pittsburgh and antithrombin with proteinases. *J Biol Chem.* 2003;278(39):37881-37887.

23. Mottonen J, Strand A, Symersky J, et al. Structural basis of latency in plasminogen activator inhibitor-1. *Nature.* 1992;355(6357):270-273.

24. Onda M, Belorgey D, Sharp LK, Lomas DA. Latent S49P neuroserpin forms polymers in the dementia familial encephalopathy with neuroserpin inclusion bodies. *J Biol Chem.* 2005;280(14):13735-13741.

25. Sui GC, Wiman B. Stability of plasminogen activator inhibitor-1: role of tyrosine221. *FEBS Lett.* 1998;423(3):319-323.

26. Tan L, Perez J, Mela M, et al. Characterising the association of latency with alpha(1)-antitrypsin polymerisation using a novel monoclonal antibody. *The international journal of biochemistry & cell biology.* 2015;58:81-91.

27. Im H, Woo MS, Hwang KY, Yu MH. Interactions causing the kinetic trap in serpin protein folding. *J Biol Chem.* 2002;277(48):46347-46354.

28. Gooptu B, Hazes B, Chang WS, et al. Inactive conformation of the serpin alpha(1)-antichymotrypsin indicates two-stage insertion of the reactive loop: implications for inhibitory function and conformational disease. *Proc Natl Acad Sci U S A.* 2000;97(1):67-72.

29. Pearce MC, Powers GA, Feil SC, Hansen G, Parker MW, Bottomley SP. Identification and characterization of a misfolded monomeric serpin formed at physiological temperature. *J Mol Biol.* 2010;403(3):459-467.

30. Horvath AJ, Irving JA, Rossjohn J, et al. The murine orthologue of human antichymotrypsin: a structural paradigm for clade A3 serpins. *J Biol Chem.* 2005;280(52):43168-43178.

31. Johnson DJ, Langdown J, Li W, Luis SA, Baglin TP, Huntington JA. Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation. *J Biol Chem.* 2006;281(46):35478-35486.

32. Lomas DA, Evans DL, Finch JT, Carrell RW. The mechanism of Z alpha 1-antitrypsin accumulation in the liver. *Nature.* 1992;357(6379):605-607.

33. Crystal RG. Alpha 1-antitrypsin deficiency, emphysema, and liver disease. Genetic basis and strategies for therapy. *J Clin Invest.* 1990;85(5):1343-1352.

34. Blanchard V, Liu X, Eigel S, et al. N-glycosylation and biological activity of recombinant human alpha1-antitrypsin expressed in a novel human neuronal cell line. *Biotechnol Bioeng.* 2011;108(9):2118-2128.

35. Sifers RN. Intracellular processing of alpha1-antitrypsin. *Proc Am Thorac Soc.* 2010;7(6):376-380.

36. Paakko P, Kirby M, du Bois RM, Gillissen A, Ferrans VJ, Crystal RG. Activated neutrophils secrete stored alpha 1-antitrypsin. *Am J Respir Crit Care Med.* 1996;154(6 Pt 1):1829-1833.

37. Cichy J, Potempa J, Travis J. Biosynthesis of alpha1-proteinase inhibitor by human lung-derived epithelial cells. *J Biol Chem.* 1997;272(13):8250-8255.

38. Cohen AB. Interrelationships between the human alveolar macrophage and alpha-1-antitrypsin. *J Clin Invest.* 1973;52(11):2793-2799.

39. Duranton J, Bieth JG. Inhibition of proteinase 3 by [alpha]1-antitrypsin in vitro predicts very fast inhibition in vivo. *Am J Respir Cell Mol Biol.* 2003;29(1):57-61.

40. Lomas DA, Mahadeva R. Alpha1-antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy. *J Clin Invest.* 2002;110(11):1585-1590.

41. Davis RL, Shrimpton AE, Holohan PD, et al. Familial dementia caused by polymerization of mutant neuroserpin. *Nature.* 1999;401(6751):376-379.

42. Gooptu B, Dickens JA, Lomas DA. The molecular and cellular pathology of alpha(1)-antitrypsin deficiency. *Trends Mol Med.* 2014;20(2):116-127.

43. Dickens JA, Lomas DA. Why has it been so difficult to prove the efficacy of alpha-1-antitrypsin replacement therapy? Insights from the study of disease pathogenesis. *Drug Des Devel Ther.* 2011;5:391-405.

44. Mills K, Mills PB, Clayton PT, Johnson AW, Whitehouse DB, Winchester BG. Identification of alpha(1)-antitrypsin variants in plasma with the use of proteomic technology. *Clin Chem.* 2001;47(11):2012-2022.

45. McCarthy C, Saldova R, O'Brien ME, et al. Increased outer arm and core fucose residues on the N-glycans of mutated alpha-1 antitrypsin

protein from alpha-1 antitrypsin deficient individuals. *J Proteome Res.* 2014;13(2):596-605.

46. Carroll TP, O'Connor CA, Floyd O, et al. The prevalence of alpha-1 antitrypsin deficiency in Ireland. *Respir Res.* 2011;12:91.

47. Lomas DA, Evans DL, Stone SR, Chang WS, Carrell RW. Effect of the Z mutation on the physical and inhibitory properties of alpha 1-antitrypsin. *Biochemistry.* 1993;32(2):500-508.

48. Knaupp AS, Bottomley SP. Structural change in beta-sheet A of Z alpha(1)-antitrypsin is responsible for accelerated polymerization and disease. *J Mol Biol.* 2011;413(4):888-898.

49. Huang X, Zheng Y, Zhang F, et al. Molecular Mechanism of Z alpha1-Antitrypsin Deficiency. *J Biol Chem.* 2016;291(30):15674-15686.

50. Gaczynska M, Karpowicz P, Stuart CE, et al. AFM Imaging Reveals Topographic Diversity of Wild Type and Z Variant Polymers of Human alpha1-Proteinase Inhibitor. *PLoS One.* 2016;11(3):e0151902.

51. Kass I, Knaupp AS, Bottomley SP, Buckle AM. Conformational properties of the disease-causing Z variant of alpha1-antitrypsin revealed by theory and experiment. *Biophys J.* 2012;102(12):2856-2865.

52. Hughes VA, Meklemburg R, Bottomley SP, Wintrode PL. The Z mutation alters the global structural dynamics of alpha1-antitrypsin. *PLoS One.* 2014;9(9):e102617.

53. Carrell RWE, D. L. I. Serpins: mobile conformations in a family of proteinase inhibitors. *Curr Opin Struct Biol.* 1992;2:438-4446.

54. Stein PE, Carrell RW. What do dysfunctional serpins tell us about molecular mobility and disease? *Nat Struct Biol.* 1995;2(2):96-113.

55. Elliott PR, Stein PE, Bilton D, Carrell RW, Lomas DA. Structural explanation for the deficiency of S alpha 1-antitrypsin. *Nat Struct Biol.* 1996;3(11):910-911.

56. Luisetti M, Seersholm N. Alpha1-antitrypsin deficiency. 1: epidemiology of alpha1-antitrypsin deficiency. *Thorax.* 2004;59(2):164-169.

57. Kroeger H, Miranda E, MacLeod I, et al. Endoplasmic reticulum-associated degradation (ERAD) and autophagy cooperate to degrade

polymerogenic mutant serpins. *J Biol Chem.* 2009;284(34):22793-22802.

58. Dickens JA, Ordonez A, Chambers JE, et al. The endoplasmic reticulum remains functionally connected by vesicular transport after its fragmentation in cells expressing Z-alpha1-antitrypsin. *FASEB J.* 2016;30(12):4083-4097.

59. Rudnick DA, Liao Y, An JK, Muglia LJ, Perlmutter DH, Teckman JH. Analyses of hepatocellular proliferation in a mouse model of alpha-1-antitrypsin deficiency. *Hepatology.* 2004;39(4):1048-1055.

60. Hidvegi T, Schmidt BZ, Hale P, Perlmutter DH. Accumulation of mutant alpha1-antitrypsin Z in the endoplasmic reticulum activates caspases-4 and -12, NFkappaB, and BAP31 but not the unfolded protein response. *J Biol Chem.* 2005;280(47):39002-39015.

61. Mahadeva R, Atkinson C, Li Z, et al. Polymers of Z alpha1-antitrypsin co-localize with neutrophils in emphysematous alveoli and are chemotactic in vivo. *Am J Pathol.* 2005;166(2):377-386.

62. Carp H, Miller F, Hoidal JR, Janoff A. Potential mechanism of emphysema: alpha 1-proteinase inhibitor recovered from lungs of cigarette smokers contains oxidized methionine and has decreased elastase inhibitory capacity. *Proc Natl Acad Sci U S A.* 1982;79(6):2041-2045.

63. Alam S, Li Z, Janciauskiene S, Mahadeva R. Oxidation of Z alpha1-antitrypsin by cigarette smoke induces polymerization: a novel mechanism of early-onset emphysema. *Am J Respir Cell Mol Biol.* 2011;45(2):261-269.

64. Haq I, Irving JA, Saleh AD, et al. Deficiency Mutations of Alpha-1 Antitrypsin. Effects on Folding, Function, and Polymerization. *Am J Respir Cell Mol Biol.* 2016;54(1):71-80.

65. Mahadeva R, Chang WS, Dafforn TR, et al. Heteropolymerization of S, I, and Z alpha1-antitrypsin and liver cirrhosis. *J Clin Invest.* 1999;103(7):999-1006.

66. Devlin GL, Chow MK, Howlett GJ, Bottomley SP. Acid Denaturation of alpha1-antitrypsin: characterization of a novel mechanism of serpin polymerization. *J Mol Biol.* 2002;324(4):859-870.

67. Mast AE, Enghild JJ, Salvesen G. Conformation of the reactive site loop of alpha 1-proteinase inhibitor probed by limited proteolysis. *Biochemistry.* 1992;31(10):2720-2728.

68. Koloczek H, Guz A, Kaszycki P. Fluorescence-detected polymerization kinetics of human alpha 1-antitrypsin. *J Protein Chem.* 1996;15(5):447-454.

69. Ekeowa UI, Freeke J, Miranda E, et al. Defining the mechanism of polymerization in the serpinopathies. *Proc Natl Acad Sci U S A.* 2010;107(40):17146-17151.

70. Elliston ELK, Lomas DA, Irving JA. In Vitro Approaches for the Assessment of Serpin Polymerization. *Methods Mol Biol.* 2018;1826:87-107.

71. James EL, Bottomley SP. The mechanism of alpha 1-antitrypsin polymerization probed by fluorescence spectroscopy. *Arch Biochem Biophys.* 1998;356(2):296-300.

72. Dafforn TR, Mahadeva R, Elliott PR, Sivasothy P, Lomas DA. A kinetic mechanism for the polymerization of alpha1-antitrypsin. *J Biol Chem.* 1999;274(14):9548-9555.

73. Chang WS, Whisstock J, Hopkins PC, Lesk AM, Carrell RW, Wardell MR. Importance of the release of strand 1C to the polymerization mechanism of inhibitory serpins. *Protein Sci.* 1997;6(1):89-98.

74. Irving JA, Haq I, Dickens JA, Faull SV, Lomas DA. Altered native stability is the dominant basis for susceptibility of alpha1-antitrypsin mutants to polymerization. *Biochem J.* 2014;460(1):103-115.

75. Zhou A, Carrell RW. Dimers initiate and propagate serine protease inhibitor polymerisation. *J Mol Biol.* 2008;375(1):36-42.

76. Miranda E, Perez J, Ekeowa UI, et al. A novel monoclonal antibody to characterize pathogenic polymers in liver disease associated with alpha1-antitrypsin deficiency. *Hepatology.* 2010;52(3):1078-1088.

77. Krishnan B, Gierasch LM. Dynamic local unfolding in the serpin alpha-1 antitrypsin provides a mechanism for loop insertion and polymerization. *Nat Struct Mol Biol.* 2011;18(2):222-226.

78. Yamasaki M, Li W, Johnson DJ, Huntington JA. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature.* 2008;455(7217):1255-1258.

79. Huntington JA, Pannu NS, Hazes B, Read RJ, Lomas DA, Carrell RW. A 2.6 A structure of a serpin polymer and implications for conformational disease. *J Mol Biol.* 1999;293(3):449-455.

80. Dunstone MA, Dai W, Whisstock JC, et al. Cleaved antitrypsin polymers at atomic resolution. *Protein Sci.* 2000;9(2):417-420.

81. Lomas DA, Finch JT, Seyama K, Nukiwa T, Carrell RW. Alpha 1-antitrypsin Siiyama (Ser53-->Phe). Further evidence for intracellular loop-sheet polymerization. *J Biol Chem.* 1993;268(21):15333-15335.

82. Nyon MP, Segu L, Cabrita LD, et al. Structural dynamics associated with intermediate formation in an archetypal conformational disease. *Structure.* 2012;20(3):504-512.

83. Carrell RW, Stein PE, Fermi G, Wardell MR. Biological implications of a 3 A structure of dimeric antithrombin. *Structure.* 1994;2(4):257-270.

84. Zhang Q, Law RH, Bottomley SP, Whisstock JC, Buckle AM. A structural basis for loop C-sheet polymerization in serpins. *J Mol Biol.* 2008;376(5):1348-1359.

85. Sharp AM, Stein PE, Pannu NS, et al. The active conformation of plasminogen activator inhibitor 1, a target for drugs to control fibrinolysis and cell adhesion. *Structure.* 1999;7(2):111-118.

86. Yamasaki M, Sendall TJ, Pearce MC, Whisstock JC, Huntington JA. Molecular basis of alpha1-antitrypsin deficiency revealed by the structure of a domain-swapped trimer. *EMBO Rep.* 2011;12(10):1011-1017.

87. Nyon MP, Prentice T, Day J, et al. An integrative approach combining ion mobility mass spectrometry, X-ray crystallography, and nuclear magnetic resonance spectroscopy to study the conformational dynamics of alpha1 -antitrypsin upon ligand binding. *Protein Sci.* 2015;24(8):1301-1312.

88. Skinner R, Chang WS, Jin L, et al. Implications for function and therapy of a 2.9 A structure of binary-complexed antithrombin. *J Mol Biol.* 1998;283(1):9-14.

89. Yamasaki M, Sendall TJ, Harris LE, Lewis GM, Huntington JA. Loop-sheet mechanism of serpin polymerization tested by reactive center loop mutations. *J Biol Chem.* 2010;285(40):30752-30758.

90. Haq I, Irving JA, Faull SV, et al. Reactive centre loop mutants of alpha-1-antitrypsin reveal position-specific effects on intermediate formation along the polymerization pathway. *Biosci Rep.* 2013;33(3).

91. Huntington JA, Whisstock JC. Molecular contortionism - on the physical limits of serpin 'loop-sheet' polymers. *Biol Chem.* 2010;391(8):973-982.

92. Bottomley SP, Hopkins PC, Whisstock JC. Alpha 1-antitrypsin polymerisation can occur by both loop A and C sheet mechanisms. *Biochem Biophys Res Commun.* 1998;251(1):1-5.

93. Dolmer K, Gettins PG. How the serpin alpha1-proteinase inhibitor folds. *J Biol Chem.* 2012;287(15):12425-12432.

94. Crowther DC, Serpell LC, Dafforn TR, Gooptu B, Lomas DA. Nucleation of alpha 1-antichymotrypsin polymerization. *Biochemistry.* 2003;42(8):2355-2363.

95. Behrens MA, Sendall TJ, Pedersen JS, Kjeldgaard M, Huntington JA, Jensen JK. The shapes of Z-alpha1-antitrypsin polymers in solution support the C-terminal domain-swap mechanism of polymerization. *Biophys J.* 2014;107(8):1905-1912.

96. Lomas DA, Hurst JR, Gooptu B. Update on alpha-1 antitrypsin deficiency: New therapies. *J Hepatol.* 2016;65(2):413-424.

97. Alpha-1 Awareness UK. Diagnosis and Treatment. 2018; http://www.alpha1.uk/links/publications/diagnosis-treatment/. Accessed 18/3/2018.

98. Lomas DA, Elliott PR, Chang WS, Wardell MR, Carrell RW. Preparation and characterization of latent alpha 1-antitrypsin. *J Biol Chem.* 1995;270(10):5282-5288.

99. Chapman KR, Burdon JG, Piitulainen E, et al. Intravenous augmentation treatment and lung density in severe alpha1 antitrypsin deficiency (RAPID): a randomised, double-blind, placebo-controlled trial. *Lancet.* 2015;386(9991):360-368.

100. Hood JM, Koep LJ, Peters RL, et al. Liver transplantation for advanced liver disease with alpha-1-antitrypsin deficiency. *N Engl J Med.* 1980;302(5):272-275.

101. Tan L, Dickens JA, Demeo DL, et al. Circulating polymers in alpha1-antitrypsin deficiency. *The European respiratory journal.* 2014;43(5):1501-1504.

102. Perlmutter DH, Brodsky JL, Balistreri WF, Trapnell BC. Molecular pathogenesis of alpha-1-antitrypsin deficiency-associated liver disease: a meeting review. *Hepatology.* 2007;45(5):1313-1323.

103. Teckman JH. Lack of effect of oral 4-phenylbutyrate on serum alpha-1-antitrypsin in patients with alpha-1-antitrypsin deficiency: a preliminary study. *J Pediatr Gastroenterol Nutr.* 2004;39(1):34-37.

104. Bouchecareilh M, Hutt DM, Szajner P, Flotte TR, Balch WE. Histone deacetylase inhibitor (HDACi) suberoylanilide hydroxamic acid (SAHA)-mediated correction of alpha1-antitrypsin deficiency. *J Biol Chem.* 2012;287(45):38265-38278.

105. Granell S, Baldini G, Mohammad S, et al. Sequestration of mutated alpha1-antitrypsin into inclusion bodies is a cell-protective mechanism to maintain endoplasmic reticulum function. *Mol Biol Cell.* 2008;19(2):572-586.

106. Wang Y, Perlmutter DH. Targeting intracellular degradation pathways for treatment of liver disease caused by alpha1-antitrypsin deficiency. *Pediatr Res.* 2014;75(1-2):133-139.

107. Hidvegi T, Ewing M, Hale P, et al. An autophagy-enhancing drug promotes degradation of mutant alpha1-antitrypsin Z and reduces hepatic fibrosis. *Science.* 2010;329(5988):229-232.

108. Kaushal S, Annamali M, Blomenkamp K, et al. Rapamycin reduces intrahepatic alpha-1-antitrypsin mutant Z protein polymers and liver injury in a mouse model. *Exp Biol Med (Maywood).* 2010;235(6):700-709.

109. Pastore N, Ballabio A, Brunetti-Pierri N. Autophagy master regulator TFEB induces clearance of toxic SERPINA1/alpha-1-antitrypsin polymers. *Autophagy.* 2013;9(7):1094-1096.

110. Chang YP, Mahadeva R, Chang WS, Lin SC, Chu YH. Small-molecule peptides inhibit Z alpha1-antitrypsin polymerization. *J Cell Mol Med.* 2009;13(8B):2304-2316.

111. Mahadeva R, Dafforn TR, Carrell RW, Lomas DA. 6-mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerization. Implications for the prevention of Z alpha(1)-antitrypsin-related cirrhosis. *J Biol Chem.* 2002;277(9):6771-6774.

112. Alam S, Wang J, Janciauskiene S, Mahadeva R. Preventing and reversing the cellular consequences of Z alpha-1 antitrypsin accumulation by targeting s4A. *J Hepatol.* 2012;57(1):116-124.

113. Chang YP, Mahadeva R, Chang WS, Shukla A, Dafforn TR, Chu YH. Identification of a 4-mer peptide inhibitor that effectively blocks the polymerization of pathogenic Z alpha1-antitrypsin. *Am J Respir Cell Mol Biol.* 2006;35(5):540-548.

114. Parfrey H, Mahadeva R, Ravenhill NA, et al. Targeting a surface cavity of alpha 1-antitrypsin to prevent conformational disease. *J Biol Chem.* 2003;278(35):33060-33066.

115. Mallya M, Phillips RL, Saldanha SA, et al. Small molecules block the polymerization of Z alpha1-antitrypsin and increase the clearance of intracellular aggregates. *J Med Chem.* 2007;50(22):5357-5363.

116. Ordonez A, Perez J, Tan L, et al. A single-chain variable fragment intrabody prevents intracellular polymerization of Z alpha1-antitrypsin while allowing its antiproteinase activity. *FASEB J.* 2015;29(6):2667-2678.

117. Motamedi-Shad N, Jagger AM, Liedtke M, et al. An antibody that prevents serpin polymerisation acts by inducing a novel allosteric behaviour. *Biochem J.* 2016;473(19):3269-3290.

118. Guo S, Booten SL, Aghajan M, et al. Antisense oligonucleotide treatment ameliorates alpha-1 antitrypsin-related liver disease in mice. *J Clin Invest.* 2014;124(1):251-261.

119. Yusa K, Rashid ST, Strick-Marchand H, et al. Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. *Nature.* 2011;478(7369):391-394.

120. Bjursell M, Porritt MJ, Ericson E, et al. Therapeutic Genome Editing With CRISPR/Cas9 in a Humanized Mouse Model Ameliorates alpha1-antitrypsin Deficiency Phenotype. *EBioMedicine.* 2018;29:104-111.

121. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. *J Struct Biol.* 1996;116(1):17-24.

122. Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol.* 2012;180(3):519-530.

123. Gest H. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes Rec R Soc Lond.* 2004;58(2):187-201.

124. Campbell ID. *Biophysical techniques.* Oxford: Oxford University Press; 2012.

125. Schermelleh L, Heintzmann R, Leonhardt H. A guide to super-resolution fluorescence microscopy. *J Cell Biol.* 2010;190(2):165-175.

126. Ruska E. Nobel lecture. The development of the electron microscope and of electron microscopy. *Biosci Rep.* 1987;7(8):607-629.

127. Merk A, Bartesaghi A, Banerjee S, et al. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell.* 2016;165(7):1698-1707.

128. Jensen G. Getting Started in Cryo-EM. 2013-2014; http://cryo-em-course.caltech.edu/.

129. Orlova EV, Saibil HR. Structural analysis of macromolecular assemblies by electron microscopy. *Chem Rev.* 2011;111(12):7710-7748.

130. Bai XC, Fernandez IS, McMullan G, Scheres SH. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife.* 2013;2:e00461.

131. Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature.* 2013;504(7478):107-112.

132. Scheres SH. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife.* 2014;3:e03665.

133. Pejchal R, Doores KJ, Walker LM, et al. A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science.* 2011;334(6059):1097-1103.

134. De Carlo S, Boisset N, Hoenger A. High-resolution single-particle 3D analysis on GroEL prepared by cryo-negative staining. *Micron.* 2008;39(7):934-943.

135. Adrian M, Dubochet J, Fuller SD, Harris JR. Cryo-negative staining. *Micron.* 1998;29(2-3):145-160.

136. Baker LA, Rubinstein JL. Radiation damage in electron cryomicroscopy. *Methods Enzymol.* 2010;481:371-388.

137. Brilot AF, Chen JZ, Cheng A, et al. Beam-induced motion of vitrified specimen on holey carbon film. *J Struct Biol.* 2012;177(3):630-637.

138. Sorzano CO, de la Fraga LG, Clackdoyle R, Carazo JM. Normalizing projection images: a study of image normalizing procedures for single particle three-dimensional electron microscopy. *Ultramicroscopy.* 2004;101(2-4):129-138.

139. Scheres SH. Semi-automated selection of cryo-EM particles in RELION-1.3. *J Struct Biol.* 2015;189(2):114-122.

140. Zhu Y, Carragher B, Glaeser RM, et al. Automatic particle selection: results of a comparative study. *J Struct Biol.* 2004;145(1-2):3-14.

141. Scheres SH. Single-particle processing in RELION-2.1. 2017; ftp://ftp.mrc-lmb.cam.ac.uk/pub/scheres/relion21_tutorial.pdf.

142. Mao Y, Wang L, Gu C, et al. Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer. *Proc Natl Acad Sci U S A.* 2013;110(30):12438-12443.

143. Henderson R. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proc Natl Acad Sci U S A.* 2013;110(45):18037-18041.

144. van Heel M, Frank J. Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy.* 1981;6(2):187-194.

145. Scheres SH, Valle M, Nunez R, et al. Maximum-likelihood multi-reference refinement for electron microscopy images. *J Mol Biol.* 2005;348(1):139-149.

146. Scheres SH, Gao H, Valle M, et al. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nature methods.* 2007;4(1):27-29.

147. Radermacher M. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. *J Electron Microsc Tech.* 1988;9(4):359-394.

148. Penczek PA, Grassucci RA, Frank J. The ribosome at improved resolution: new techniques for merging and orientation refinement in

3D cryo-electron microscopy of biological particles. *Ultramicroscopy.* 1994;53(3):251-270.

149. Van Heel M. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy.* 1987;21(2):111-123.

150. Grigorieff N. FREALIGN: high-resolution refinement of single particle structures. *J Struct Biol.* 2007;157(1):117-125.

151. van Heel M, Orlova EV, Harauz G, et al. Angular Reconstitution in Three-Dimensional Electron Microscopy: Historical and Theoretical Aspects. *Scanning Microscopy* 1997;11:195.

152. Penczek PA. Fundamentals of three-dimensional reconstruction from projections. *Methods Enzymol.* 2010;482:1-33.

153. Herman GT, Lent A. Iterative reconstruction algorithms. *Comput Biol Med.* 1976;6(4):273-294.

154. Crowther RA, DeRosier DJ, Klug A. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proc R Soc London* 1970;317:319.

155. Friedrich H, de Jongh PE, Verkleij AJ, de Jong KP. Electron tomography for heterogeneous catalysts and related nanostructured materials. *Chem Rev.* 2009;109(5):1613-1629.

156. Drulyte I, Johnson RM, Hesketh EL, et al. Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallogr D Struct Biol.* 2018;74(Pt 6):560-571.

157. van Heel M, Schatz M. Fourier shell correlation threshold criteria. *J Struct Biol.* 2005;151(3):250-262.

158. Rosenthal PB, Crowther DC, Henderson R. An Objective Criterion for Resolution Assessment in Single-particle Electron Microscopy. *J Mol Biol.* 2003;333:743-745.

159. Kucukelbir A, Sigworth FJ, Tagare HD. Quantifying the local resolution of cryo-EM density maps. *Nature methods.* 2014;11(1):63-65.

160. Winn MD, Ballard CC, Cowtan KD, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011;67(Pt 4):235-242.

161. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004;60(Pt 12 Pt 1):2126-2132.

162. van Heel M, Keegstra W. IMAGIC: A fast, flexible and friendly image analysis software system. *Ultramicroscopy.* 1981;7(2):113-129.

163. *The PyMOL Molecular Graphics System* [computer program]. 2002.

164. Chaudhury S, Lyskov S, Gray JJ. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics.* 2010;26(5):689-691.

165. Vidarsson G, Dekkers G, Rispens T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol.* 2014;5:520.

166. Alberts B. *Molecular biology of the cell.* 5th ed. New York ; Abingdon: Garland Science; 2008.

167. Schroeder HW, Jr., Cavacini L. Structure and function of immunoglobulins. *J Allergy Clin Immunol.* 2010;125(2 Suppl 2):S41-52.

168. Sela-Culang I, Alon S, Ofran Y. A systematic comparison of free and bound antibodies reveals binding-related conformational changes. *J Immunol.* 2012;189(10):4890-4899.

169. An Z. *Therapeutic monoclonal antibodies : from bench to clinic.* Hoboken, NJ: Wiley-Blackwell; 2009.

170. Hu X, Kang S, Lefort C, Kim M, Jin MM. Combinatorial libraries against libraries for selecting neoepitope activation-specific antibodies. *Proc Natl Acad Sci U S A.* 2010;107(14):6252-6257.

171. Stamenova M, Kehayov I, Kyurkchiev S, Despodova T. Production and application of monoclonal antibody specific for human chorionic gonadotropin. *Folia Biol (Praha).* 1990;36(2):81-90.

172. Bain B, Brazil M. Adalimumab. *Nat Rev Drug Discov.* 2003;2(9):693-694.

173. Cai HH. Therapeutic monoclonal antibodies approved by FDA in 2017. *Journal of Immunology.* 2018;6(3):82-84.

174. Chudasama V, Maruani A, Caddick S. Recent advances in the construction of antibody-drug conjugates. *Nat Chem.* 2016;8(2):114-119.

175. Beck A, Goetsch L, Dumontet C, Corvaia N. Strategies and challenges for the next generation of antibody-drug conjugates. *Nat Rev Drug Discov.* 2017;16(5):315-337.

176. Greenfield EA. *Antibodies : a laboratory manual.* Second edition. ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2014.

177. Zhang C. Hybridoma technology for the generation of monoclonal antibodies. *Methods Mol Biol.* 2012;901:117-135.

178. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature.* 1975;256(5517):495-497.

179. Yokoyama WM. Production of monoclonal antibody supernatant and ascites fluid. *Curr Protoc Mol Biol.* 2008;Chapter 11:Unit 11 10.

180. Polyclonal and Monoclonal Antibody Production. *Microbiology* https://courses.lumenlearning.com/microbiology/chapter/polyclonal-and-monoclonal-antibody-production/.

181. Irving JA, Miranda E, Haq I, et al. An antibody raised against a pathogenic serpin variant induces mutant-like behaviour in the wild-type protein. *Biochem J.* 2015;468(1):99-108.

182. Goldsteins G, Persson H, Andersson K, et al. Exposure of cryptic epitopes on transthyretin only in amyloid and in amyloidogenic mutants. *Proc Natl Acad Sci U S A.* 1999;96(6):3108-3113.

183. Vaikath NN, Majbour NK, Paleologou KE, et al. Generation and characterization of novel conformation-specific monoclonal antibodies for alpha-synuclein pathology. *Neurobiol Dis.* 2015;79:81-99.

184. Hatami A, Albay R, 3rd, Monjazeb S, Milton S, Glabe C. Monoclonal antibodies against Abeta42 fibrils distinguish multiple aggregation state polymorphisms in vitro and in Alzheimer disease brain. *J Biol Chem.* 2014;289(46):32131-32143.

185. Bodker JS, Wind T, Jensen JK, Hansen M, Pedersen KE, Andreasen PA. Mapping of the epitope of a monoclonal antibody protecting plasminogen activator inhibitor-1 against inactivating agents. *Eur J Biochem.* 2003;270(8):1672-1679.

186. Wind T, Jensen MA, Andreasen PA. Epitope mapping for four monoclonal antibodies against human plasminogen activator inhibitor

type-1: implications for antibody-mediated PAI-1-neutralization and vitronectin-binding. *Eur J Biochem.* 2001;268(4):1095-1106.

187. Berry CN, Lunven C, Lechaire I, Girardot C, O'Connor SE. Antithrombotic activity of a monoclonal antibody inducing the substrate form of plasminogen activator inhibitor type 1 in rat models of venous and arterial thrombosis. *British journal of pharmacology.* 1998;125(1):29-34.

188. Bijnens AP, Gils A, Stassen JM, et al. The distal hinge of the reactive site loop and its proximity: a target to modulate plasminogen activator inhibitor-1 activity. *J Biol Chem.* 2001;276(48):44912-44918.

189. Wallmark A, Alm R, Eriksson S. Monoclonal antibody specific for the mutant PiZ alpha 1-antitrypsin and its application in an ELISA procedure for identification of PiZ gene carriers. *Proc Natl Acad Sci U S A.* 1984;81(18):5690-5693.

190. Janciauskiene S, Dominaitiene R, Sternby NH, Piitulainen E, Eriksson S. Detection of circulating and endothelial cell polymers of Z and wild type alpha 1-antitrypsin by a monoclonal antibody. *J Biol Chem.* 2002;277(29):26540-26546.

191. Miranda E, Ferrarotti I, Berardelli R, et al. The pathological Trento variant of alpha-1-antitrypsin (E75V) shows nonclassical behaviour during polymerization. *FEBS J.* 2017;284(13):2110-2126.

192. Gooptu B, Miranda E, Nobeli I, et al. Crystallographic and cellular characterisation of two mechanisms stabilising the native fold of alpha1-antitrypsin: implications for disease and drug design. *J Mol Biol.* 2009;387(4):857-868.

193. Fra A, Cosmi F, Ordonez A, et al. Polymers of Z alpha1-antitrypsin are secreted in cell models of disease. *The European respiratory journal.* 2016;47(3):1005-1009.

194. Zhu XJ, Chan SK. The use of monoclonal antibodies to distinguish several chemically modified forms of human alpha 1-proteinase inhibitor. *Biochem J.* 1987;246(1):19-23.

195. Ueda M, Mashiba S, Uchida K. Evaluation of oxidized alpha-1-antitrypsin in blood as an oxidative stress marker using anti-oxidative alpha1-AT monoclonal antibody. *Clin Chim Acta.* 2002;317(1-2):125-131.

196. Subramaniyam D, Glader P, von Wachenfeldt K, Burneckiene J, Stevens T, Janciauskiene S. C-36 peptide, a degradation product of alpha1-antitrypsin, modulates human monocyte activation through LPS signaling pathways. *The international journal of biochemistry & cell biology.* 2006;38(4):563-575.

197. Janciauskiene S, Zelvyte I, Jansson L, Stevens T. Divergent effects of alpha1-antitrypsin on neutrophil activation, in vitro. *Biochem Biophys Res Commun.* 2004;315(2):288-296.

198. Ducruix A, Giegé R. *Crystallization of nucleic acids and proteins : a practical approach.* 2nd ed. Oxford: Oxford University Press; 1999.

199. Rupp B. *Biomolecular crystallography : principles, practice, and application to structural biology.* New York: Garland Science; 2010.

200. Garman EF, Owen RL. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2006;62(Pt 1):32-47.

201. Scapin G, Yang X, Prosise WW, et al. Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab. *Nat Struct Mol Biol.* 2015;22(12):953-958.

202. Terry WD, Matthews BW, Davies DR. Crystallographic studies of a human immunoglobulin. *Nature.* 1968;220(5164):239-241.

203. Guddat LW, Herron JN, Edmundson AB. Three-dimensional structure of a human immunoglobulin with a hinge deletion. *Proc Natl Acad Sci U S A.* 1993;90(9):4271-4275.

204. Perez J, Garrido O, Cifuentes M, et al. Bovine Reissner's fiber (RF) and the central canal of the spinal cord: an immunocytochemical study using a set of monoclonal antibodies against the RF-glycoproteins. *Cell Tissue Res.* 1996;286(1):33-42.

205. Wang Z, Raifu M, Howard M, et al. Universal PCR amplification of mouse immunoglobulin gene variable regions: the design of degenerate primers and an assessment of the effect of DNA polymerase 3' to 5' exonuclease activity. *J Immunol Methods.* 2000;233(1-2):167-177.

206. Whisstock JC, Skinner R, Carrell RW, Lesk AM. Conformational changes in serpins: I. The native and cleaved conformations of alpha(1)-antitrypsin. *J Mol Biol.* 2000;296(2):685-699.

207. Zhou A, Carrell RW, Huntington JA. The serpin inhibitory mechanism is critically dependent on the length of the reactive center loop. *J Biol Chem.* 2001;276(29):27541-27547.

208. Irving JA, Pike RN, Dai W, et al. Evidence that serpin architecture intrinsically supports papain-like cysteine protease inhibition: engineering alpha(1)-antitrypsin to inhibit cathepsin proteases. *Biochemistry.* 2002;41(15):4998-5004.

209. Plotnick MI, Samakur M, Wang ZM, et al. Heterogeneity in serpin-protease complexes as demonstrated by differences in the mechanism of complex breakdown. *Biochemistry.* 2002;41(1):334-342.

210. Schousboe SL, Egelund R, Kirkegaard T, Preissner KT, Rodenburg KW, Andreasen PA. Vitronectin and substitution of a beta-strand 5A lysine residue potentiate activity-neutralization of PA inhibitor-1 by monoclonal antibodies against alpha-helix F. *Thromb Haemost.* 2000;83(5):742-751.

211. Lijnen HR, Arza B, Van Hoef B, Collen D, Declerck PJ. Inactivation of plasminogen activator inhibitor-1 by specific proteolysis with stromelysin-1 (MMP-3). *J Biol Chem.* 2000;275(48):37645-37650.

212. Stura EAF, G. G.; Wilson, I. A. Crystallization of Antibodies and Antibody-Antigen complexes. *Immunomethods.* 1993;3:164-179.

213. Bloch Y, Bouchareychas L, Merceron R, et al. Structural Activation of Pro-inflammatory Human Cytokine IL-23 by Cognate IL-23 Receptor Enables Recruitment of the Shared Receptor IL-12Rbeta1. *Immunity.* 2018;48(1):45-58 e46.

214. Borst AJ, James ZM, Zagotta WN, et al. The Therapeutic Antibody LM609 Selectively Inhibits Ligand Binding to Human alphaVbeta3 Integrin via Steric Hindrance. *Structure.* 2017;25(11):1732-1739 e1735.

215. Talavera A, Mackenzie J, Garrido G, et al. Structure of the Fab fragment of the anti-murine EGFR antibody 7A7 and exploration of its receptor binding site. *Mol Immunol.* 2011;48(12-13):1578-1585.

216. Winter G, McAuley KE. Automated data collection for macromolecular crystallography. *Methods.* 2011;55(1):81-93.

217. Zwart PH, Grosse-Kunstleve, R. W., Adams, P. D. Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. *CCP4 Newsletter.* 2005;43(Contribution 7).

218. Lescar J, Stouracova R, Riottot MM, et al. Three-dimensional structure of an Fab-peptide complex: structural basis of HIV-1 protease inhibition by a monoclonal antibody. *J Mol Biol.* 1997;267(5):1207-1222.

219. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007;40(Pt 4):658-674.

220. Skubak P, Murshudov GN, Pannu NS. Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr D Biol Crystallogr.* 2004;60(Pt 12 Pt 1):2196-2201.

221. Adams PD, Afonine PV, Bunkoczi G, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 2):213-221.

222. Weik M, Ravelli RB, Kryger G, et al. Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc Natl Acad Sci U S A.* 2000;97(2):623-628.

223. Lovell SC, Davis IW, Arendall WB, 3rd, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins.* 2003;50(3):437-450.

224. Dr. Andrew C.R. Martin's Group U. Antibodies. 2018, July 30; http://www.bioinf.org.uk/abs/ - cdrdef.

225. Al-Lazikani B, Lesk AM, Chothia C. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol.* 1997;273(4):927-948.

226. Sela-Culang I, Kunik V, Ofran Y. The structural basis of antibody-antigen recognition. *Front Immunol.* 2013;4:302.

227. Ramaraj T, Angel T, Dratz EA, Jesaitis AJ, Mumey B. Antigen-antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. *Biochim Biophys Acta.* 2012;1824(3):520-532.

228. Jeffrey GA. *An introduction to hydrogen bonding.* New York ; Oxford: Oxford University Press; 1997.

229. Davies DR, Cohen GH. Interactions of protein antigens with antibodies. *Proc Natl Acad Sci U S A.* 1996;93(1):7-12.

230. Patschull AO, Segu L, Nyon MP, et al. Therapeutic target-site variability in alpha1-antitrypsin characterized at high resolution. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2011;67(Pt 12):1492-1497.

231. Cabrita LD, Dai W, Bottomley SP. Different conformational changes within the F-helix occur during serpin folding, polymerization, and proteinase inhibition. *Biochemistry.* 2004;43(30):9834-9839.

232. Andersen OJ, Risor MW, Poulsen EC, et al. Reactive Center Loop Insertion in alpha-1-Antitrypsin Captured by Accelerated Molecular Dynamics Simulation. *Biochemistry.* 2017;56(4):634-646.

233. Jin L, Abrahams JP, Skinner R, Petitou M, Pike RN, Carrell RW. The anticoagulant activation of antithrombin by heparin. *Proc Natl Acad Sci U S A.* 1997;94(26):14683-14688.

234. Belzar KJ, Zhou A, Carrell RW, Gettins PG, Huntington JA. Helix D elongation and allosteric activation of antithrombin. *J Biol Chem.* 2002;277(10):8551-8558.

235. Tsumoto KC, J. MM. Antigen-Antibody binding. *Encyclopedia of Life Sciences*: John Wiley & Sons; 2016.

236. Couture ML, Heath CA. Relationship between loss of heavy chains and the appearance of nonproducing hybridomas. *Biotechnol Bioeng.* 1995;47(2):270-275.

237. Westerwoudt RJ. Factors affecting production of monoclonal antibodies. *Methods Enzymol.* 1986;121:3-18.

238. Kousted TM, Skjoedt K, Petersen SV, et al. Three monoclonal antibodies against the serpin protease nexin-1 prevent protease translocation. *Thromb Haemost.* 2014;111(1):29-40.

239. Eyal E, Gerzon S, Potapov V, Edelman M, Sobolev V. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *J Mol Biol.* 2005;351(2):431-442.

240. Qi X, Loiseau F, Chan WL, et al. Allosteric modulation of hormone release from thyroxine and corticosteroid-binding globulins. *J Biol Chem.* 2011;286(18):16163-16173.

241. Grunberg R, Nilges M, Leckner J. Flexibility and conformational entropy in protein-protein binding. *Structure.* 2006;14(4):683-693.

242. Galant NJ, Bugyei-Twum A, Rakhit R, et al. Corrigendum: Substoichiometric inhibition of transthyretin misfolding by immune-targeting sparsely populated misfolding intermediates: a potential diagnostic and therapeutic for TTR amyloidoses. *Sci Rep.* 2016;6:27679.

243. Bijnens AP, Ngo TH, Gils A, et al. Elucidation of the binding regions of PAI-1 neutralizing antibodies using chimeric variants of human and rat PAI-1. *Thromb Haemost.* 2001;85(5):866-874.

244. Perrie AM, MacGregor, I. R., Booth, N. A. Definition of Epitopes Within Plasminogen Activator Inhibitor Type-1 (PAI-1) Using Multiple Peptide Synthesis. *Fibrinolysis.* 1993;7:257-263.

245. Lawrence DA, Olson ST, Muhammad S, et al. Partitioning of serpin-proteinase reactions between stable inhibition and substrate cleavage is regulated by the rate of serpin reactive center loop insertion into beta-sheet A. *J Biol Chem.* 2000;275(8):5839-5844.

246. Bottomley SP, Lawrenson ID, Tew D, Dai W, Whisstock JC, Pike RN. The role of strand 1 of the C beta-sheet in the structure and function of alpha(1)-antitrypsin. *Protein Sci.* 2001;10(12):2518-2524.

247. Lane DA, Olds RJ, Conard J, et al. Pleiotropic effects of antithrombin strand 1C substitution mutations. *J Clin Invest.* 1992;90(6):2422-2433.

248. Scheidig AJ, Hynes TR, Pelletier LA, Wells JA, Kossiakoff AA. Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzheimer's amyloid beta-protein precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): engineering of inhibitors with altered specificities. *Protein Sci.* 1997;6(9):1806-1824.

249. Hansen G, Gielen-Haertwig H, Reinemer P, Schomburg D, Harrenga A, Niefind K. Unexpected active-site flexibility in the structure of human neutrophil elastase in complex with a new dihydropyrimidone inhibitor. *J Mol Biol.* 2011;409(5):681-691.

250. Faull SV. *Characterising Alpha-1-Antitrypsin Polymers Using Biochemical and Biophysical Techniques*, University of Cambridge; 2015.

251. Tang G, Peng L, Baldwin PR, et al. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol.* 2007;157(1):38-46.

252. Kremer JR, Mastronarde DN, McIntosh JR. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol.* 1996;116(1):71-76.

253. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol.* 2003;142(3):334-347.

254. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol.* 1999;128(1):82-97.

255. Harpaz Y, Gerstein M, Chothia C. Volume changes on protein folding. *Structure.* 1994;2(7):641-649.

256. Bohne-Lang A, von der Lieth CW. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.* 2005;33(Web Server issue):W214-219.

257. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods.* 2012;9(7):676-682.

258. Scheres SH. A Bayesian view on cryo-EM structure determination. *J Mol Biol.* 2012;415(2):406-418.

259. Scheres SH. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods Enzymol.* 2016;579:125-157.

260. Fra AM, Gooptu B, Ferrarotti I, et al. Three new alpha1-antitrypsin deficiency variants help to define a C-terminal region regulating conformational change and polymerization. *PLoS One.* 2012;7(6):e38405.

261. Aebi U, ten Heggeler B, Onorato L, Kistler J, Showe MK. New method for localizing proteins in periodic structures: Fab fragment labeling combined with image processing of electron micrographs. *Proc Natl Acad Sci U S A.* 1977;74(12):5514-5518.

262. Wu S, Avila-Sakar A, Kim J, et al. Fabs enable single particle cryoEM studies of small proteins. *Structure.* 2012;20(4):582-592.

263. Kim J, Wu S, Tomasiak TM, et al. Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. *Nature.* 2015;517(7534):396-400.

264. Griffin L, Lawson A. Antibody fragments as tools in crystallography. *Clin Exp Immunol.* 2011;165(3):285-291.

265. De Carlo S, Harris JR. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron.* 2011;42(2):117-131.

266. Negative Staining Techniques. *Bio-Imaging* 2004; http://web.path.ox.ac.uk/~bioimaging/bitm/instructions_and_information/em/neg_stain.pdf.

267. Thompson RF, Walker M, Siebert CA, Muench SP, Ranson NA. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods.* 2016;100:3-15.

268. Tsutsui Y, Kuri B, Sengupta T, Wintrode PL. The structural basis of serpin polymerization studied by hydrogen/deuterium exchange and mass spectrometry. *J Biol Chem.* 2008;283(45):30804-30811.

269. Berg NO, Eriksson S. Liver disease in adults with alpha-1 -antitrypsin deficiency. *N Engl J Med.* 1972;287(25):1264-1267.

270. Sharp HL, Bridges RA, Krivit W, Freier EF. Cirrhosis associated with alpha-1-antitrypsin deficiency: a previously unrecognized inherited disorder. *J Lab Clin Med.* 1969;73(6):934-939.

271. Gordon HW, Dixon J, Rogers JC, Mittman C, Lieberman J. Alpha 1 -antitrypsin (A1AT) accumulation in livers of emphysematous patients with A1AT deficiency. *Hum Pathol.* 1972;3(3):361-370.

272. Fitzpatrick AWP, Falcon B, He S, et al. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature.* 2017;547(7662):185-190.

273. Falcon B, Zhang W, Murzin AG, et al. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature.* 2018;561(7721):137-140.

274. Guerrero-Ferreira R, Taylor NM, Mona D, et al. Cryo-EM structure of alpha-synuclein fibrils. *Elife.* 2018;7.

275. Bauerlein FJB, Saha I, Mishra A, et al. In Situ Architecture and Cellular Interactions of PolyQ Inclusions. *Cell.* 2017;171(1):179-187 e110.

276. Guedes-Dias P, Holzbaur ELF. Huntingtin Fibrils Poke Membranes. *Cell.* 2017;171(1):32-33.

277. Chaban Y, Stead JA, Ryzhenkova K, et al. Structural basis for DNA strand separation by a hexameric replicative helicase. *Nucleic Acids Res.* 2015;43(17):8551-8563.

278. Fischer N, Neumann P, Konevega AL, et al. Structure of the E. coli ribosome-EF-Tu complex at <3 A resolution by Cs-corrected cryo-EM. *Nature.* 2015;520(7548):567-570.

279. Bartesaghi A, Merk A, Banerjee S, et al. 2.2 A resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor. *Science.* 2015;348(6239):1147-1151.

280. Eriksson S, Larsson C. Purification and partial characterization of pas-positive inclusion bodies from the liver in alpha 1-antitrypsin deficiency. *N Engl J Med.* 1975;292(4):176-180.

281. Bolmer S, Kleinerman J. Isolation and characterization of alpha 1-antitrypsin in PAS-positive hepatic granules from rats with experimental alpha 1-antitrypsin deficiency. *Am J Pathol.* 1986;123(2):377-389.

282. Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol.* 2005;152(1):36-51.

283. Kobs G. Endo H Application: Monitoring Protein Trafficking. 2012; https://www.promegaconnections.com/endo-h-application-monitoring-protein-trafficking/11003ma-2/.

284. Ermantraut E, Wohlfart, K., Tichelaar, W. Perforated support foils with pre-defined hole size, shape and arrangement. *Ultramicroscopy.* 1998;74(1):75-81.

285. Russo CJ, Passmore LA. Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. *J Struct Biol.* 2016;193(1):33-44.

286. Bathurst IC, Travis J, George PM, Carrell RW. Structural and functional characterization of the abnormal Z alpha 1-antitrypsin isolated from human liver. *FEBS Lett.* 1984;177(2):179-183.

287. Chatani E, Lee YH, Yagi H, Yoshimura Y, Naiki H, Goto Y. Ultrasonication-dependent production and breakdown lead to minimum-sized amyloid fibrils. *Proc Natl Acad Sci U S A.* 2009;106(27):11119-11124.

288. Hercz A, Katona E, Cutz E, Wilson JR, Barton M. alpha1-Antitrypsin: the presence of excess mannose in the Z variant isolated from liver. *Science.* 1978;201(4362):1229-1232.

289. Dawwas MF, Davies SE, Griffiths WJ, Lomas DA, Alexander GJ. Prevalence and risk factors for liver involvement in individuals with PiZZ-related lung disease. *Am J Respir Crit Care Med.* 2013;187(5):502-508.

290. Passmore LA, Russo CJ. Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol.* 2016;579:51-86.

291. Ellgaard L, McCaul N, Chatsisvili A, Braakman I. Co- and Post-Translational Protein Folding in the ER. *Traffic.* 2016;17(6):615-638.

292. Jonic S. Cryo-electron Microscopy Analysis of Structurally Heterogeneous Macromolecular Complexes. *Comput Struct Biotechnol J.* 2016;14:385-390.

293. Cheng Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell.* 2015;161(3):450-457.

294. Vinothkumar KR, Henderson R. Single particle electron cryomicroscopy: trends, issues and future perspective. *Q Rev Biophys.* 2016;49:e13.

295. Nakane T, Kimanius D, Lindahl E, Scheres SH. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *Elife.* 2018;7.

296. Tskhovrebova L, Trinick J. Flexibility and extensibility in the titin molecule: analysis of electron microscope data. *J Mol Biol.* 2001;310(4):755-771.

297. Lee JH, Ozorowski G, Ward AB. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science.* 2016;351(6277):1043-1048.

298. Sotriffer CA, Rode BM, Varga JM, Liedl KR. Elbow flexibility and ligand-induced domain rearrangements in antibody Fab NC6.8: large effects of a small hapten. *Biophys J.* 2000;79(2):614-628.

299. Khoshouei M, Radjainia M, Baumeister W, Danev R. Cryo-EM structure of haemoglobin at 3.2 A determined with the Volta phase plate. *Nat Commun.* 2017;8:16099.

300. Fernandez-Leiro R, Conrad J, Scheres SH, Lamers MH. cryo-EM structures of the E. coli replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and tau. *Elife.* 2015;4.

301. Palovcak E, Wang F, Zheng SQ, et al. A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *J Struct Biol.* 2018;204(1):80-84.

302. Danev R, Baumeister W. Cryo-EM single particle analysis with the Volta phase plate. *Elife.* 2016;5.

303. Engh R, Lobermann H, Schneider M, Wiegand G, Huber R, Laurell CB. The S variant of human alpha 1-antitrypsin, structure and implications for function and metabolism. *Protein Eng.* 1989;2(6):407-415.

304. Laffranchi M, Berardelli R, Ronzoni R, Lomas DA, Fra A. Heteropolymerization of alpha-1-antitrypsin mutants in cell models mimicking heterozygosity. *Hum Mol Genet.* 2018;27(10):1785-1793.